This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED LINDED THE DATENT COOPERATION

(51) International Patent Classification ⁶ : A01K 67/027, 67/033, C12N 15/00		12	(11)	International Publication Number:	WO 98/28971	
AUIK 0//02/, 0	77035, C12N 15/00	A2	(43)	International Publication Date:	9 July 1998 (09.07.98)	
(21) International Appli	cation Number: PCT/US	97/238	19 (81) Designated States: AL, AM, AT, A	U, AZ, BA, BB, BG, BR,	
(22) International Filing	Date: 31 December 1997 (31.12.9	7)	BY, CA, CH, CN, CU, CZ, DE, GH, GM, GW, HU, ID, IL, IS, J LC, LK, LR, LS, LT, LU, LV, I	P, KE, KG, KP, KR, KZ,	
(30) Priority Data: 08/778,645	3 January 1997 (03.01.97)	· [JS	MX, NO, NZ, PL, PT, RO, RU, TJ, TM, TR, TT, UA, UG, US, U patent (GH, GM, KE, LS, MW, S	SD, SE, SG, SI, SK, SL, JZ, VN, YU, ZW, ARIPO D, SZ, UG, ZW), Eurasian	
60/080,222	3 January 1997 (03.01.97)	l	JS	patent (AM, AZ, BY, KG, KZ, M	D. RU, TJ, TM), Europea	

(71) Applicant (for all designated States except US): UNIVER-SITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventor; and

- (75) Inventor/Applicant (for US only): LINK, Christopher [US/US]; 1837 Spruce Street, Boulder, CO 80302 (US).
- (74) Agents: MONROY, Gladys, H. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: BETA-AMYLOID TOXICITY

(57) Abstract

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes in vivo. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, Caenorhabditis elegans for visualization of the toxic effects of β -amyloid accumulation in vivo.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	. TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL.	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

BETA-AMYLOID TOXICITY

FIELD OF THE INVENTION

The present invention is in the field of medicine and molecular biology. In particular, the invention provides transgenic non-human animals in which expression of a reporter gene is induced by a toxic transgene. The invention also provides methods and compositions for *in vivo* visualization of the toxicity associated with toxic proteins. In addition, the present invention provides methods and compositions for the production and use of transgenic non-human animals for screening candidate drugs to assess and treat toxicity.

10

15

20

25

30

5

BACKGROUND OF THE INVENTION

Investigation of the molecular mechanisms of various diseases has suggested that many pathologies are caused by expression of proteins that are immediately or gradually deleterious to the cells where they are expressed. Among these diseases are Alzheimer's disease, prion diseases, Huntington's disease, and amyotrophic lateral sclerosis. Several of these diseases are thought to result from aberrant folding of proteins, which results in the accumulation of toxic proteins or protein aggregates.

Alzheimer's disease is associated with the majority of dementia cases in the United States, with an estimated 2 million people afflicted with the disease, and a mortality rate of approximately 100,000 people per year (See, R.W.P. Cutler, "Degenerative and Hereditary Diseases," in Medicine, Scientific American, New York, (1988), pages 11 (IV):1-13; and R. Katzman (1986) N. Engl. J. Med. 314:964). It has been estimated that the total cost for nursing home care alone of Alzheimer's patients exceeds \$13 billion/year (See, M.M. Heckler (1985) Am. Psychol. 40:1240). According to the Centers for Disease Control (CDC), mortality due to Alzheimer's disease in the United States increased 10-fold between 1979 and 1987 (0.4 per 100,000 to 4.2 per 100,000) (See, "Reported death rate for Alzheimer's is up tenfold since 1979," (1990) Clin. Psychiatr. News 18:21).

Patients suffering from Alzheimer's disease typically suffer progressive memory deficit, progressive decline in cognitive functions, anxiety, depression, visuospatial and speech deficits, delusions, personality changes, motor skill deterioration, loss of verbal

ability, and incontinence. Eventually, patients are completely incapacitated and disoriented, requiring total care. The course of Alzheimer's disease ranges from less than three years to over 20 years before death occurs. However, in typical cases, it progresses at a fairly constant rate, with an average duration of 6 to 10 years.

5

10

15

Alzheimer's disease is not the only cause of dementia. Indeed, there are over fifty recognized causes of dementia. As some causes of dementia are amenable to treatment, differential diagnosis of patients suffering dementia is particularly important. Because at least 20% of clinically diagnosed patients were found at autopsy to have had conditions other than Alzheimer's disease, the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and Alzheimer's Disease and Related Disorders Association (ADRDA) refined the clinical diagnostic criteria for Alzheimer's disease (*See*, M.A. Jenike, "Psychiatry," *in Medicine*, Scientific American, New York, [1991], pages 13 (V):1-5). Based on these criteria, the diagnosis of Alzheimer's disease may be "definite" (*i.e.*, requiring examination of brain tissue), "probable" (*i.e.*, patients have deficits in two or more areas of cognition, insidious onset of disease, progressive worsening of memory and other cognitive functions, and normal consciousness levels), or "possible" (*i.e.*, patients meet the criteria for probable Alzheimer's disease, but exhibit variations in the disease course or have a systemic illness that is sufficient to cause dementia, but is not considered to be the cause of the dementia).

20

The numerous varieties of dementia, and variations in patient presentations, often make diagnosis problematic. Thus, the NINCDS/ADRDA criteria are very detailed, and necessitate the thorough examination of patients with suspected Alzheimer's disease. Currently, the only way to obtain a definite diagnosis is by post-mortem histological examination of brain tissue for the presence of senile plaques.

25

30

The histopathological lesions of Alzheimer's disease include neuritic or senile plaques, neurofibrillary degeneration, and granulovacuolar neuronal degeneration. The senile plaques usually contain a core of insoluble, amyloidic extracellular material (" β -amyloid) surrounded by a halo of neurofibrillary tangles and dystrophic neurons. The primary protein component of the amyloidic core of senile plaques is a 4.2 kd amyloid β peptide (often referred to as " $A\beta$ "). A number of other proteins have also been identified

as components of senile plaques, including anti-chymotrypsin and apolipoprotein E. The major component of the neurofibrillary tangles is an abnormally phosphorylated microtubule-associated protein referred to as "tau." The β-amyloid protein in senile plaques is a small glycoprotein (*i.e.*, a 39-43 amino acid protein, derived from the larger amyloid precursor protein), that has been detected in non-neural tissues (*e.g.*, skin, subcutaneous tissue, and intestines) and blood vessels of Alzheimer's disease patients (*See*, Cutler, *supra*). Deposits of amyloid may be detected by their ability to bind specific dyes, such as Congo red or thioflavin S. Thus, it has been further hypothesized that detection of this protein may serve as a potential diagnostic aid in the assessment of patients suffering from dementia.

5

10

15

20

25

30

Alzheimer's disease usually arises spontaneously, although genetics play a role in development of the disease. For example, the histopathologic lesions of Alzheimer's disease have been regularly observed in the brains of older patients dying of Down's syndrome (*See*, Cutler, *supra*). It is known that the β-amyloid gene resides on chromosome 21. As Down's patients have an extra copy of this chromosome, there is an increased expression of this chromosome in these patients. It has been hypothesized that increased expression of the β-amyloid gene may regulate the formation of amyloid plaques ("senile plaques") in these patients. In addition, four loci were recently identified as playing a role in the genetic susceptibility of Alzheimer's disease (*See*, Pericak-Vance and Haines (1995) *Trends Genet.* 11:504).

Other potential risk factors for the development of Alzheimer's disease include environmental factors (e.g., head trauma, smoking, and exposure to heavy metals), sociological factors (e.g., depression and educational level), biological factors (e.g., increasing age and hyperthyroidism), and a family history of Alzheimer's disease, Down's syndrome, or Parkinson's disease (See, Pericak-Vance and Haines, supra). Nonetheless, despite recent advances, the exact etiology and pathogenesis of Alzheimer's disease remain largely unknown.

In addition to the problems associated with diagnosing Alzheimer's disease, improvements are needed in the area of treatment. Many agents have been tested for their ability to treat the cognitive decline associated with Alzheimer's disease. For example,

various cholinergic enhancers (e.g., choline and lecithin) have been tested. Unfortunately, cholinergic precursors have been shown to be not useful, although some drugs that stimulate cholinergic transmission may be helpful in some patients. One example is physostigmine, a compound that prevents the synaptic breakdown of acetylcholine. However, the overall clinical effect of this drug has not been as dramatic as initially hoped (See, M.A. Jenike, "Psychiatry," in Medicine, Scientific American, New York, (1991), pages 13 (V):1-5).

Other drugs, such as tetrahydroaminoacridine (THA or tacrine), a centrally acting anti-cholinesterase, have been tested. In a large multi-center trial of THA, liver enzyme abnormalities were reported and the preliminary results indicated that, at least at low dosages, THA is not an effective treatment of Alzheimer's disease (Jenike, *supra*). Ergoloid mesylates (Hydergine) is an extremely safe compound, and remains the most commonly prescribed drug for patients with Alzheimer's disease. However, the overall effects of the drug are at best minimal.

15

5

10

The mechanisms of toxicity due to accumulation of β -peptide are currently the subject of much investigation, and no definitive causes of toxicity have yet been established. (See, Benzi and Moretti (1995) *Neurobiology of Aging*, **16**:661-674). Because of this uncertainty, efforts to develop mechanism-based treatment regimens have not been possible. Compounds presently in use to treat Alzheimer's disease only serve to alleviate the systemic effects associated with the disease.

20

25

30

In order to study disease mechanisms and genetic-based phenomena, animals in which a foreign gene has been inserted have been described by various researchers. International Patent Application WO 96/03034 describes insertion of retroviral vectors into fish, in order to produce fish with desirable traits or to study development. Various transgenic animal models for Alzheimer's disease are described in International Patent Applications WO 93/14200, WO 93/02189, WO 94/12627, WO 94/23049, and European Patent Publication EP 653154. Typically these transgenic animals are mice or other mammals; however, β-amyloid peptide has been expressed in the nematode *Caenorhabditis elegans* (Link (1995) *Proc. Natl. Acad. Sci.* 92: 9368). The use of *C. elegans* for investigation of mutant forms of the *C. elegans* genes mec-4 and deg-1, which

cause neurodegeneration, is described in U.S. Patent No. 5,196,333. Drawbacks to mammalian animals as model systems are the relatively long generation time, which makes mammals less desirable for high-throughput screening of potential pharmaceuticals, and the difficulty in studying the molecular processes of interest without sacrificing the animal for cell and tissue analysis.

Reporter genes are genes that encode proteins or other compounds that can be detected by a variety of methods, and which "report" the occurrence of successful introduction and expression of gene sequences. β-galactosidase and luciferase are examples of such reporter genes. Recently, the gene for green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* has been described in U.S. Patent No. 5,491,084. This gene can provide a method for indicating expression of a gene via fluorescent detection of GFP. The drawback to using GFP in mammalian systems is as described above, that is, the need to sacrifice the mammal in order to analyze the cells and tissues of interest. This drawback is especially severe in the context of large-scale screening of potential therapeutic compounds, and U.S. Patent No. 5,491,084 does not describe a system suitable for high-throughput screening of pharmaceuticals for activity against toxic proteins expressed in cells.

Thus, methods and animal systems are needed to screen drugs quickly and inexpensively for their effects on proteins and other substances associated with Alzheimer's disease. Convenient methods and animal systems for screening drugs for other neurodegenerative diseases, such as prion diseases, Huntington's disease, and amyotrophic lateral sclerosis, are also desirable.

SUMMARY OF THE INVENTION

25

30

5

10

15

20

The present invention provides methods and compositions useful for the production and use of dually transgenic animals, in particular *Caenorhabditis elegans*. In addition, in one embodiment, the present invention provides transparent animals that express a reporter gene inducible by a toxic transgene. In a preferred embodiment, the toxic transgene encodes β -peptide. Although it is not intended that the present invention be limited to any particular reporter, in a preferred embodiment, the reporter gene is green fluorescence

protein (GFP). It is contemplated that the effects or product of the reporter gene be observable (*i.e.*, the reporter gene is expressed). For example, in embodiments in which GFP is the reporter, its presence is detected using fluorescence microscopy. In addition, in these embodiments, the animals do not need to be sacrificed in order to observe the expression of the reporter gene (*i.e.*, the detection may be accomplished on living animals). If other reporters are used, other detection methods may be necessary. For example, lacZ expression may be detected by exposing the tissues of the animal to the substrate for the gene (*i.e.*, β -galactoside), and observing for the presence of blue dye in the tissues. However, this method requires that the animal be sacrificed in order to observe the expression of the reporter gene.

5

10

15

20

25

In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In a preferred embodiment, the transparent animal is selected from the class Nematoda. Although it is not intended that the transparent animal of the present invention be limited to any specific animal, in a particularly preferred embodiment, the transparent animal is *Caenorhabditis elegans*.

In an alternate embodiment the cells of the transparent animal display toxicity resulting from the accumulation of β -peptide within the cells. In one preferred embodiment, the genome of the transparent animal comprises SEQ ID NO:5 and SEQ ID NO:8.

5

10

The present invention also provides methods for producing dually transgenic non-human animal comprising: providing: a first and second non-human animal; a first transgene comprising β -peptide; and a second transgene, comprising a reporter; introducing the first transgene into the genome of the first non-human animal to produce a first transgenic animal, and introducing the second transgene into the genome of the second non-human animal to produce a second transgenic animal; and mating the first transgenic animal with the second transgenic animal to produce a dually transgenic animal, wherein the β -peptide and the reporter are expressed.

In one embodiment of the methods of the present invention the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis* elegans.

20

15

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

25

The present invention also comprises methods for testing compounds for anti-toxic effects, comprising providing a dually transgenic non-human animal expressing a toxic transgene and a reporter; a composition comprising a test compound in a form suitable for administration such that the compound is bioavailable in the cells of the animal; and administering the test compound to the non-human animal. The toxic transgene induces

30

5

10

15

20

25

30

expression of the reporter gene. In a preferred embodiment, the toxic transgene is β -peptide and the reporter gene is GFP. In one embodiment, the method further comprises the step of measuring a reduction or increase in the expression of the reporter by the dually transgenic non-human animal and thereby identifying a compound as therapeutic. In a particularly preferred embodiment of the methods, the compounds inactivate the β -peptide expressed by the dually transgenic animal.

In one embodiment of the methods for testing compounds for β-peptide toxicity, the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis elegans*. In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is *rol-6(su-1006)*, which produces effects described herein.

In one alternative embodiment, one or more test compounds are tested for their ability to counter the toxic effects of transgene product. It is also contemplated that the test compounds will be tested for their ability to prevent the expression of the toxic transgene, for example, β -peptide.

In alternative embodiments, dually transgenic animals in which the expression of, or effects of transgene toxicity, are reduced or eliminated by the test compounds, are mated. The progeny of these matings are also then tested for the effects of test compounds on the expression of, or effects of toxicity. In yet another embodiment, the progeny of these matings are used in other assay systems for the identification of therapeutic compounds.

In yet another alternative embodiment, the dually transgenic animals of the present invention are used to identify methods suitable for the diagnostic testing of Alzheimer's disease or pathology due to other diseases. Thus, the dually transgenic animals of the present invention are used to develop assays suitable for use in humans or animal models of Alzheimer's disease.

DESCRIPTION OF THE FIGURES

Figure 1 is a diagram illustrating the construction of dual transgenic animals expressing β-amyloid peptide and an hsp/GFP reporter.

Figure 2 shows the DNA sequence (SEQ ID NO:5) and restriction map for pCL12. Figure 3 is a graphic map of pCL12.

Figure 4 shows the DNA (SEQ ID NO:6), and amino acid sequence (SEQ ID NO:7) of pCL12 from nucleotide 1071 through 1253 (*i.e.*, the β -(1-42) nucleic acid and amino acid sequence).

Figure 5 shows the DNA sequence (SEQ ID NO:8) and restriction map for pGFP-TT.

Figure 6 is a graphic map of pGFP-TT.

5

15

20

25

30

Figure 7 shows the DNA sequence (SEQ ID NO:9) and restriction map for pCL25. Figure 8 is a graphic map of pCL25.

Figure 9 shows the DNA sequence (SEQ ID NO:10) of *rol*-6. Figure 10 shows the amino acid sequence (SEQ ID NO:11) of *rol*-6.

DESCRIPTION OF THE INVENTION

The present invention provides methods and compositions useful for the production and use of transgenic animals. The methods and transgenic animals of the invention also provide an efficient and effective system for screening drug effective in ameloriating the effects of toxic gene products.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, cell biology and recombinant DNA, which are within the skill of the art. *See, e.g.*, Sambrook, Fritsch, and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, (F.M. Ausubel et al. eds., 1987); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.); PCR 2: A PRACTICAL APPROACH (M.J. McPherson, B.D. Hames and G.R. Taylor eds., 1995); ANIMAL CELL CULTURE (R.I.

Freshney. Ed., 1987); and ANTIBODIES: A LABORATORY MANUAL (Harlow et al. eds., 1987).

Animals

5

10

15

The methods and systems described herein can be practiced with any non-human animal. In a preferred embodiment, the animal used is transparent for at least part of its life cycle, for example, oocytes (*Xenopus* or others); larvae; pupae, fish (zebrafish or others); and nematodes, such as *C. elegans*. *C. elegans* is particularly preferred because it is transparent throughout its life and has been extensively studied as a model to determine developmentally regulated gene expression, as well as pattern formation. The visibility of all of the animals' cells throughout their life cycles, has led to the complete determination of the cell lineages, and a detailed description of the morphogenic changes that occur during *C. elegans* development. Rapid methods for introduction of cloned DNA into the *C. elegans* germ line has provided means to study molecular function and expression *in vivo*, as the effects of the gene expression are usually readily visible in the transparent worms.

Reporter Molecules

20

The methods and compositions of the present invention make use of reporter genes to monitor gene activity. Visible and quantifiable reporter genes are known and described in the art. Successfully used reporter molecules in gene fusion vectors in studies with *C. elegans* include the *E. coli lacZ* coding region (*See e.g.*, A. Fire *et al.*, (1990) *Gene* **93**:189-198), and the *Aequora victoria gfp* coding region (*See e.g.*, M. Chalfie *et al.*, (1994) *Science* **263**:802-805; D.C. Prasher *et al.*, (1992) *Gene* **111**:229-233; and Genbank

25

Accession #M62654), which produces an intrinsically fluorescent protein. Although various fusion expression vectors have been used and reported in the literature, problems have been encountered (See e.g., A. Fire, "Fire Lab Vector Kit"--June 1995). For example, with lacZ, there have been many reports in which it was not possible to correlate transgene expression patterns with physiological expectations, or only weak correlations have been

30

possible. Ectopic expression is often frequently seen with short promoters, and occurs

most prominently in the gut and pharynx of the affected animals. This may be due to weak promoter and/or enhancer signals in the vectors. Expression pattern deficits of lacZ fusions have been classified into three groups. In the first, transgenes are expressed in the correct tissue, but mosaic expression is observed (*i.e.*, only a subset of the cells stain during the detection methods). This has been observed even with integrated high-copy transgenes. In the second group, expression in a single tissue or cell population is not seen with the transgene. Finally, there has been the failure of transgene fusion constructs to show expression in the pre-12 cell embryo, or in any embryonic, larval. or adult germ line.

The *gfp* fusion vectors were developed as an alternative to the *lacZ* markers, but have been available for a much shorter time (*See e.g.*, Chalfie *et al.*, (1994), *supra*; U.S. Patent No. 5,491,084). Initial reports indicate that the fluorescence pattern appears to be more restricted than that exhibited by equivalent *lacZ* fusion. In addition, there are still problems associated with the germline expression of the transgenes and ectopic expression. Indeed, in some cases, the problems with ectopic expression have been exacerbated by these "improvements."

The present invention overcomes many of these problems described in the art. The transparent animals and dually transgenic animals described herein provide easily visible, stably expressed systems in which the toxic transgene and reporter gene are expressed similarly to endogenous chromosomal gene expression. These animals provide an excellent system for screening compounds having effects on the toxic transgene. The present invention, therefore, provides model systems for the study of human diseases and methods of identifying therapeutic compounds using these animal systems.

Inducible Promoters

25

30

20

5

10

15

As discussed above, the reporter genes are operably linked to an inducible promoter. The promoter is induced by the toxic transgene, for example by the gene product of the toxic transgene. It is also contemplated that the toxic transgene can act to induce the promoter indirectly, for example by disrupting other cellular proteins or functions. Suitable inducible promoters are available and can be readily determined by those skilled in the art. Non-limiting examples of promoters which are induced by

"stresses" include the metallothionein gene promoter (*mtl*-1 or *mtl*-2; *e.g.*, Genbank Accession #M92910, #M11794, #X00504, and #X00953), and the *C. elegans* amyloid precursor protein (APP) homolog *apl*-1 gene promoter (*See e.g.*, Daigle and Li, Proc. Natl. Acad., Sci., 90:12045-12049 [1993]). Other potential promoters include those from other inducible heat-shock genes; at least one of the known *C. elegans hsp*70 genes is strongly heat-inducible. It is also contemplated that promoters from genes known to be upregulated under stress conditions in other systems (*e.g.*, superoxide dismutase, catalase, glutathione reductase, etc.) may also be useful.

5

10

15

20

25

In one embodiment, as described in the Examples below, a strong muscle-specific promoter was used to express a potentially secretable form of the β -peptide, so as to generate significant extracellular levels of β -peptide. This was accomplished in order to mimic the situation that may exist in the human brain, and allow observations of cell-external neurotoxicity. The present inventor has also demonstrated that the promoter must be chosen so that enough β -peptide is produced to cause physiological effects, but not so much to kill the animals.

The unc-54/ β -(1-42) minigene was constructed with a modified signal sequence that has been previously shown to allow secretion of a her-1 protein product that is ectopically expressed in muscle cells (M.D. Perry $et\ al.$, (1993) $Genes\ Dev.$, 7:216-228). The development of $C.\ elegans$ transgenic for expression of β -amyloid was previously described by the inventor (Link, (1995) $Proc.\ Natl.\ Acad.\ Sci.$, 92:9368-9572).

However, in early experiments, it was observed that β -peptide deposits were not convincingly detected outside of the muscle cells, when tested with the antibodies described in Example 4, below. While an understanding of the mechanism is not necessary for the practice of the present invention, it is apparent that the majority of β -peptide expressed by the transgenic worms is retained in the muscle cells and is responsible for the pathology observed in the muscle cells.

It is also contemplated that β -peptide expression will be directed to other tissues, through utilization of appropriate promoters. For example, it is contemplated that animals expressing β -peptide in the intestine may be particularly useful to analyze compounds such

as drugs for their effect on β -peptide, as these cells readily take up exogenous compounds administered orally.

Transgenes

5

The transgene element of the present invention can be any sequence which is able to induce the promoter operably linked to the reporter gene. As noted, the transgene is "toxic" in the sense that it disrupts cellular function in some way. Preferably, the transgene encodes a protein that is toxic to the host cells and/or organism in that it causes deleterious effects to the host, for instance interfering with the hosts ability to survive and/or grow. The toxic transgenes encode proteins that disrupt cellular function directly (e.g., the gene product is toxic) or indirectly (e.g., the sequence of the transgene disrupts cellular function by some mechanism other than its gene product). Toxic proteins are distinguishable from simple chemical toxins (e.g., heavy metals and the like) by their antigenicity and higher molecular weight.

15

10

In one embodiment, the toxic transgene comprises a gene encoding for an amyloidic protein, for example beta amyloid peptide, prion protein variants, transthyretin variants, gelsolin variants, cystatin variants, lysozyme variants and the like. In another embodiment, the transgene encodes a protein containing polyglutamine resulting from triplet-repeat expansion such as huntingtin (a protein that has been implicated in Huntington's Chorea), ataxin-1 or ataxin-2. Alternatively, proteins associated with inherent amyelotopic lateral sclerosis (ALS) for example, superoxide dismutase 1 variants and over-expressed neurofilament protein, can be used. It will be understood that the transgene can encode for an entire toxic protein or, alternatively, a functional (i.e. toxic) fragment.

25

30

20

Transgenic Animals

In one aspect, the present invention includes a dually transgenic non-human animal. In one embodiment, the development of these dually transgenic animals involves the production of two *C. elegans* lines with a single transgene in each line, designated as CL2005 and CL2070. *C. elegans* was chosen due to the observation that the toxic effect of

β-amyloid has been reported to occur in these animals, in addition to humans. Line CL2005 exhibited muscle-specific expression of human β peptide, while line CL2070 exhibited stress-inducible expression of GFP. GFP was incorporated into the present invention as it provides a marker (*i.e.*, "reporter") that is readily visible in living worms. Thus, dually transgenic animals both express β-peptide in their muscle cells and under appropriate conditions for observation (*i.e.*, fluorescence microscopy), exhibit green fluorescence in these cells.

Mating of these lines resulted in the production of dually transgenic animals, in which the presence of β -amyloid was easy to detect, due to the fact that the dually transgenic animals expressing β -amyloid glowed green. Thus, the present invention provides an easily detectable method for the specific expression of β -peptide in dually transgenic animals. This ease of detection provides great advantages for the development of methods to analyze the effects of β -amyloid *in vivo*. In addition, because the dually transgenic animals are not killed in order to detect the presence of β -amyloid, as would be required if other reporters (*e.g., lacZ* or luciferase) were used, the animals may be useful for screening compounds for their effects on β -amyloid in an *in vivo* situation. Thus, the use of an reporter which is detectable *in vivo* provides significant advantages over currently available methods.

Other Advantages of Transgenic C. elegans

5

10

15

20

25

30

The present invention also includes methods of screening compounds for their ability to prevent or inhibit toxicity due to the toxic transgene. In these drug screening embodiments, the methods and animals described above can be used to identify potential therapies. For example, in assays to develop compounds to block the expression of β -amyloid or counter its toxic effects, transgenic worms (*e.g.*, green fluorescing worms with muscle-specific β -peptide expression), would be observed for their loss of fluorescence. In assays where the compound is effective in countering the expression of β -peptide or effects of β -peptide, the animals would no longer glow green. It is easy to visually screen and quantify (using commercially available equipment) for animals that have lost the ability to glow, and then use them to detail the interaction between the compound and the toxic

5

10

15

20

25

30

transgene. The present invention provides a fast and efficient screening system, for example by using commercially available equipment for assaying multiple compounds at once.

Thus, the transgenic animals of the present invention expressing a detectable reporter gene triggered by a toxic transgene provide animal models for human diseases. For instance, a reporter gene (such as GFP) induced by expression of β-peptide provides an animal model for human Alzheimer's disease. It is also not intended that the present invention be limited to animals expressing β-peptide in conjunction with GFP. It is contemplated that expression of other proteins associated with dementia and/or Alzheimer's or other diseases would also find use in conjunction with GFP in *C. elegans* as well. For example, it is contemplated that proteins (including mutated versions of proteins) such as the presenilins associated with Alzheimer's disease (*See e.g.*, J. Marx (1996) *Science* 274:1838-1840) will be used in conjunction with GFP and full-length APP in dually transgenic *C. elegans*. Other toxic proteins which provide animal models for other human diseases are described herein.

These transgenic animals also provide a means to screen compounds for their ability to decrease or eliminate the toxicity. Examples of compounds suitable for testing using the transgenic animals of the present invention include, but are not limited to, Congo Red, tumor necrosis factor (TNF), estrogen, tacrine (9-amino-1,2,3,4-tetrahydroacridine), dihydroepiandrosterone (DHEA), compounds that inhibit ApoE4, and others, commercially available from suppliers such as Sigma. Compounds to be tested for antitoxic activity are administered to the same number of dually transgenic animals (e.g., generated using the methods described in the Examples, below) from the control group and the treatment group, and the presence or absence of reporter gene used as a measure of efficacy.

The compounds being tested can be administered using any suitable route (e.g., oral, parenteral, controlled-release transdermal methods, and implants, etc.). In one preferred route, the compounds to be tested are suspended in the growth media provided to the worms. Generally speaking, the route of administration will depend on the stability of the compound, the susceptibility of the compound to "first pass" metabolism, the

5

10

15

20

25

30

concentration needed to achieve a therapeutic effect, and the like. Following initial screening, a compound that appears promising (*i.e.*, which increases the number of worms which display reduced β -peptide toxicity) is further evaluated by administering various concentrations of the compound to additional transgenic animals in order to determine an approximate therapeutic dosing range.

Another screening method involves the crossing of the transgenic worms of the present invention with other transgenic worms. The animals are observed after treatment, in the presence and absence of the test compound(s), with the effects on the toxic transgene being gauged either by crude survival or the presence/absence of the reporter gene. It is also contemplated that the methods of the present invention be modified so as to provide means to analyze disease-related proteins believed to have dominant toxic effects, by substituting appropriate sequences for the β-peptide sequences used as described in the Examples. For example, proteins such as transthyretin (known to be associated with familial amyloid polyneuropathy; See e.g., Christmason et al., (1991) FEBS 281:177-180; Genbank Accession #D00096), and variant superoxide dismutase (known to be associated with familial amyotrophic lateral sclerosis [Lou Gehrig's disease]), prion proteins, A4 amyloid protein (See e.g., Ponte et al., (1988) Nature 331:525-527; Salbaum et al., U.S. Patent No. 5,151,508, herein incorporated by reference), APP (See e.g., Kitaguchi et al., (1988) Nature 331:530-532; Sata et al., EP Appln. 94117512.7; Scott et al., WO 9412627; Wadsworth et al., WO 9314200; Gearhart et al., WO 9423049; and Neve et al., WO 9302189, all of which are herein incorporated by reference), other amyloidic proteins (e.g., variant lysozymes and amylin peptide), and other proteins associated with neuronal degeneration (See e.g., U.S. Patent No. 5,196,333), will be used in the methods of the present invention by substituting the appropriate nucleic acid sequences encoding the protein of interest for the β-peptide described in Example 1. It is also contemplated that animals transgenic for dual proteins (e.g., β -peptide in combination with another protein, such as transthyretin) will be used. Thus, the present invention provides the means to analyze the effects of numerous genes and proteins in vivo.

Furthermore, the transgenic animals of the present invention provide distinct advantages over other transgenic animals currently used to analyze diseases such as

Alzheimer's disease. Due to the short gestation period of *C. elegans*, transgenic animals can be produced much more rapidly than when mammals, such as mice are used. For example, transgenic mice overexpressing a 695 amino acid isoform of β -amyloid precursor did not show learning and memory impairment until they were 9-10 months of age (Hsiao *et al.*, Science 274:99-102 [1996]). In contrast, the transgenic animals of the present invention express β -peptide and may be manipulated at a very early age, even in embryonic stages.

5

10

15

20

25

30

In addition, because the transgenic animals are easy to select based on the presence of the green fluorescence, invasive procedures such as surgery, necessary to analyze the effects of the toxic transgene in other animals (e.g., mice) are avoided.

Finally, the apparent toxicity of β-peptide in transgenic animals was found to be temperature-dependent. Transgenic animals maintained at 25.5°C were significantly more sick than those maintained at 16°C. The animals maintained at the higher temperature became paralyzed more quickly, failed to eat or grow, failed to show normal egg-laying, and many died before reaching adulthood. However, wild-type animals grow well at either of these temperatures. The CL2005 parental line was temperature-sensitive for viability, as it was not possible to propagate this line at the elevated temperature. In the CL2070 parental line, 25.5°C maintenance was not sufficient to induce the hsp/GFP construct. The hsp/GFP response was similarly temperature-dependent. It was difficult to detect GFP when the dual transgenic animals were raised at 16°C, but GFP induction was dramatic when the animals were raised at 25.5°C.

These temperature dependency observations were exploited by the propagation of animals at 16°C, and then upshifting them to 25.5°C, when GFP induction was desired. This effect also has potential benefits in the analysis of compounds, as the animals can be pre-incubated in the presence of drugs or other compounds at the lower temperature before the upshift to the higher temperature, in order to ensure that the presumed protective effect of the drug was in place, prior to the strong induction of β -peptide toxicity. Importantly, these assays may be completed within one day and are suitable for rapid methods (*e.g.*, the use of a microtiter format and a plate fluorimeter), so that literally thousands of compounds may be tested simultaneously.

In sum, the present invention provides methods and compositions useful as animal models for disease, as well as providing methods and compositions for disease therapy and prevention. The animal testing may be supplemented and confirmed by testing on human subjects. However, the transgenic animals of the present invention allow the testing of a large number of compounds, both various methods, including those known in the art.

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below.

The term "non-human animal" includes vertebrates such as rodents, arthropods, insects (e.g., Diptera), fish (e.g., zebrafish), non-human primates, ovines, bovines, ruminants, lagomorphs, porcines, caprines, equines, canines, felines, aves, etc. The non-human animal can be used at any stage in its development, for example oocyte, fetal, larval, pupal stages or the like. Preferred non-human animals are those that are transparent, such as certain nematodes, oocytes, larvae and fish.

15

20

5

10

As used herein, the term "transparent" is used in references to animals through which light will be transmitted. However, it is not intended that the amount of light transmittance be limited to any particular amount. For example, an animal is transparent so long as at least some light may be transmited through its body. Non-limiting examples of transparent animals are larval stages of some animals (e.g., flies) and oocytes (e.g., Xenopus oocytes). Particularly preferred non-human animals are selected from the nematodes (i.e., any animal in the Class Nematoda), most preferably Caenorhabditis elegans. However, it is contemplated that other transparent animals, such as zebrafish will be useful in the present invention. For example, transgenic zebrafish have been produced (See e.g., Lin et al., PCT Publ. WO9603034; incorporated herein by reference).

25

30

The "non-human animals having a genetically engineered genotype" of the invention are preferably produced by experimental manipulation of the genome of the germline of the non-human animal. These genetically engineered non-human animals may be produced by several methods including the introduction of a "transgene" comprising nucleic acid (usually DNA) into an embryonal target cell or integration into a chromosome of the somatic and/or germ line cells of a non-human animal by way of human

intervention, such as by the methods described herein. The process by which a DNA molecule becomes stably incorporated into another genome is referred to as "stable integration." Non-human animals which contain a transgene are referred to as "transgenic non-human animals". A transgenic animal is an animal whose genome has been altered by the introduction of one or more transgenes.

5

10

15

20

25

30

The term "transgene" as used herein refers to a foreign gene that is placed into an organism by introducing the foreign gene into gonadal cells, embryonic cells, newly fertilized eggs, or early embryos. The term "foreign gene" refers to any nucleic acid (e.g., gene sequence) which is introduced into the genome of an animal by experimental manipulations and may include gene sequences found in that animal so long as the introduced gene contains some modification relative to the naturally-occurring gene. A preferred foreign gene is the β -amyloid gene (e.g., β -peptide), or fragments thereof. A "toxic transgene" refers to a foreign gene which disrupts cellular function in some way. \Box For example, the toxic transgene may produce a gene product (i.e., protein) which is toxic to the cell or organism. Alternatively, the transgene may disrupt other cellular proteins, or act as a DNA binding site without producing a protein.

As used herein, the term "toxicity" refers to the production of toxic effects by any compound or substance. For example, while it is not necessary to the understanding and use of the present invention, β -peptide toxicity may arise due to the accumulation of β -peptide molecules. There are numerous manifestations of toxicity that may occur. For example, the toxicity associated with β -peptide may be manifested as negative effects on muscle cells of dually transgenic animals.

The terms "promoter element, " "promoter," or "promoter sequence" as used herein, refer to a DNA sequence that is located at the 5' end of (*i.e.*, precedes) a gene in a DNA polymer and provides a site for initiation of the transcription of the gene into mRNA. An "inducible promoter" is a promoter that is triggered by certain signals within the cell, for instance binding of a transcription factor, stress, heat or the like.

The term "reporter gene" as used herein refers to genes that encode proteins or other compounds that can be detected by a variety of methods. These genes "report" the occurrence of successful introduction and expression of sequences such as transgenes.

Non-limiting examples of reporter genes include antibiotic resistance genes, genes encoding enzymes and genes encoding other detectable proteins. Expression of the reporter gene is detected using methods known in the art. In a preferred embodiment, the reporter gene is a GFP gene, although lacZ β -galactosidase gene, or any other reporter system may be used to detect the successful production of transgenic animals. In particularly preferred embodiments, the reporter is a compound or protein which may be present or expressed within living animals. That is, it is not necessary to sacrifice the animal in order to detect the presence of the reporter. The type of the reporter gene system used is not critical to the invention, and it is contemplated that any system suitable for use with the transgenic animals of the present invention will be used.

5

10

15

20

25

30

As used herein, the terms "dual trangenic" and "dually transgenic" refer to animals or cells in which more than one transgene have been introduced. For example, the term is used in reference to cells which contain the sequences encoding β -peptide, and sequences encoding a reporter (e.g., GFP). However, it is not intended that the number of transgenes in the dually transgenic animals of the present invention be limited to two. For example, the transgenic animals may also contain another one or more marker genes (e.g., rol-6 sequences), in addition to β -peptide and reporter sequences. In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

The transgenic animals of the present invention are preferentially generated by introduction of the targeting vectors into gonad cells. Transgenes can be efficiently introduced into the cells by DNA transfection using a variety of methods known to the art, including electroporation, calcium phosphate co-precipitation, protoplast or spheroplast fusion, lipofection, and DEAE-dextran-mediated transfection. Transgenes may also be introduced into cells by retrovirus-mediated transduction or by micro-injection. In one preferred embodiment, the transgenes are injected into gonads of *C. elegans* as described by Mello *et al.* (1991) *EMBO J.*, **10**:3959-3970. Alternative methods for the generation of transgenic animals containing an altered gene are known to the art. For example,

5

10

15

20

25

embryonal cells at various developmental stages can be used to introduce transgenes for the production of transgenic animals. Different methods are used depending on the stage of development of the embryonal cell.

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting related sequences include oligolabeling, nick translation, and end-labeling or PCR amplification using a labeled nucleotide. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. The term may also be used in reference to proteins. For example, a variety of protocols which employ polyclonal or monoclonal antibodies specific for the β-peptide protein product are known in the art (*See*, the Examples). These antibodies can be used as markers for the expression of proteins.

Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS).

The term "gene" refers to a DNA sequence that comprises coding sequences necessary for the production of a polypeptide or precursor. In some instances, a gene can also include control sequences. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired enzymatic activity is retained.

The term "gene of interest" refers to any gene, the manipulation of which may be deemed desirable for any reason, by one of ordinary skill in the art.

The term "wild-type" refers to a gene or gene product which has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designed the "normal" or "wild-type" form of the gene. In contrast, the term "modified" or "mutant" refers to a gene or gene product which displays modifications in sequence and or functional properties (*i.e.*, altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are

identified by the fact that they have altered characteristics when compared to the wild-type gene or gene product.

The terms "targeting vector" or "targeting construct" refer to oligonucleotide sequences comprising a gene of interest flanked on either side by regulatory sequences. Preferably, the targeting vector is capable of homologous recombination such that the gene of interest is integrated by recombination.

5

10

15

20

25

30

As used herein, the terms "vector" and "vehicle" are used interchangeably in reference to nucleic acid molecules that transfer DNA segment(s) from one cell to another.

The terms "expression vector" or "expression cassette" as used herein, refers to a recombinant DNA molecule containing a desired coding sequence and appropriate nucleic acid sequences necessary for the expression of the operably linked coding sequence in a particular host organism. Nucleic acid sequences necessary for expression in prokaryotes usually include a promoter, an operator (optional), and a ribosome binding site, often along with other sequences. Eukaryotic cells are known to utilize promoters, enhancers, and termination and polyadenylation signals.

The terms "in operable combination," "in operable order," and "operably linked" as used herein, refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of amino acid sequences in such a manner so that a functional protein is produced.

The term "tissue-type specific" as it applies to a promoter, refers to a promoter that is capable of selectively directing expression of a gene in a specific tissue. Similarly, the use of a tissue-specific promoter in the method of the present invention does not require absolute specificity. In general, the requisite specificity is found where a plurality (or, more preferably, a majority) of cells in one tissue type express a gene of interest, while virtually all (e.g., greater than 80%, and preferably greater than 90%, and more preferably greater than 95%), of the cells in other tissue types do not. In one embodiment of the present invention, the strong muscle promoter *unc*54 was used (*See*, Example 1).

As used herein the term "portion" when in reference to a gene refers to fragments of that gene. The fragments may range in size from a few nucleotides to the entire gene

sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

As used herein, the term "substantially purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. An "isolated polynucleotide" is therefore a substantially purified polynucleotide.

5

10

15

20

25

30

As used herein, the term "polymerase chain reaction" ("PCR") refers to the method of K.B. Mullis disclosed in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,965,188, all of which are hereby incorporated by reference, which describe a method for increasing the concentration of a segment of a target sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the target sequence consists of introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective strands of the double stranded target sequence. To effect amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the target molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing and polymerase extension can be repeated many times (i.e., denaturation, annealing and extension constitute one "cycle"; there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired target sequence. The length of the amplified segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of the repeating aspect of the process, the method is referred to as the "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the target sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified."

. Œ

With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (e.g., hybridization

with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; and/or incorporation of ³²P-labeled deoxynucleotide triphosphates, such as dCTP or dATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide sequence can be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications. Amplified target sequences may be used to obtain segments of DNA (*e.g.*, genes) for the construction of targeting vectors, transgenes, etc.

5

10

15

20

25

30

As used herein, the terms "PCR product" and "amplification product" refer to the resultant mixture of compounds after two or more cycles of the PCR steps of denaturation, annealing and extension are complete. These terms encompass the case where there has been amplification of one or more segments of one or more target sequences.

The term "Central Nervous System" refers to the "spinal cord" and the "brain." The spinal cord comprises white areas and grey areas. The grey area contains nerve cell bodies, whereas the white area is essentially comprised of myelinated nerves. The brain, which is also known as the encephalon, is that portion of the cerebrospinal axis which is contained in the cavity of the cranium. The brain comprises the two cerebral hemispheres, the interbrain, the mid-brain, the pons Varolli and cerebellum, and the medulla oblongata. The two hemispheres together with the parts derived from the thalamencephalon form the forebrain. The two cerebral hemispheres are separated by the longitudinal fissure and also comprise the bylvian fissure, the fissure of Rolando, and the parieto-occipital fissure. The lobes on the external surface of the brain comprise the frontal lobe, the parietal lobe, the occipital lobe, and the temporal lobe. Placed along the middle line of the brain are, among others, the rostrum and peduncles of corpus callosum, lamina cinera, optic commissure and the pituitary body. On each side of the middle line lies the frontal lobe, olfactory lobe and the hemisphere of cerebellum.

The terms "neuron," "neural cell," and "nerve cell" are used interchangeably to refer to a cell which is located in the nervous system. Nerve cells are composed of the nerve cell body (perikaryon), one or more dendrites, and an axon. Neurons can be classified according to the number of processes originating from the cell body. Thus,

unipolar neurons have a single process, bipolar neurons have one axon and one dendrite, while multipolar neurons (which are the most common) comprise more than two processes. The term "neuron" comprises cholinergic neurons and sensory neurons. As used herein, the term "cholinergic neuron" means a neuron in the Central Nervous System (CNS) and in the Peripheral Nervous System (PNS) whose neurotransmitter is acetylcholine. As used herein, the term "sensory neuron" includes a neuron which is responsive to environmental cues (e.g., temperature and movement) from, for example, the skin, muscle and joints of a mammal.

5

10

15

20

25

30

The term "nerve" refers to two or more neurons arranged in linear sequence such that the axon of one neuron establishes a structural and functional link with the dendrite of a second neuron to form a "synapse."

The term "compound" refers to any chemical entity, pharmaceutical, drug, and the like that can be used to treat or prevent a disease, illness, sickness, or disorder of bodily function. Compounds comprise both known and potential therapeutic compounds. A compound can be determined to be therapeutic by testing using the testing methods of the present invention (*i.e.*, a "test compound"). A "known therapeutic compound" refers to a therapeutic compound that has been shown (*e.g.*, through animal trials or prior experience with administration to humans) to be effective in such treatment or prevention. In other words, a known therapeutic compound is not limited to a compound efficacious in the treatment or prevention of neural related disorders.

A compound is said to be "in a form suitable for administration such that the compound is bioavailable in the blood of the animal" when the compound may be administered to an animal by any desired route (e.g., oral, intravenous, subcutaneous, intramuscular, etc.) and the compound or its active metabolites appears in the blood of the animal in an active form. Administration of a compound to a pregnant animal may result in delivery of bioavailable compound to the embryonic progeny of the animal.

The "wild-type β -amyloid" or " β -peptide" gene and gene product refers to the nucleotide and amino acid sequences provided in SEQ ID NOS:6 and 7, respectively. Those skilled in the art will be well aware that certain modifications of SEQ ID NOS:6 and 7 can be made which will not interfere with the production of a polypeptide having an

activity indistinguishable from that of the wild-type β-amyloid; the present invention specifically contemplates these variant forms of β-amyloid. A "variant" of the β-peptide is defined as an amino acid sequence that differs by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions (i.e., additions), or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNAStar software.

5

10

15

20

25

30

The term "an oligonucleotide sequence comprising at least a portion of a human β -amyloid gene" refers to a polynucleotide sequence (*i.e.*, a nucleic acid sequence) containing a nucleotide sequence derived from a human β -amyloid gene. This sequence may encode a portion or all of the β -amyloid protein; alternatively, this sequence may contain non-coding regions derived from the β -amyloid gene or a combination of coding and non-coding regions. The oligonucleotide may be RNA or DNA and may be of genomic or synthetic origin.

As used herein the term "portion" when in reference to a gene refers to fragments of that gene. The fragments may range in size from 10 nucleotides to the entire gene sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

An animal whose genome comprises a "heterologous marker gene" is an animal whose genome contains a marker gene not naturally found in the animal's genome. In one preferred embodiment of the present invention, the heterologous marker gene is a mutant collagen gene, such as the rol-6(su-1006) gene. However, it is intended that other marker genes will be used with success in the present invention, including other mutant collagen genes, as well as other marker genes commonly known to those in the art.

As used herein, the term "diagnostic assay" refers to methods for the diagnosis of disease, illness, and/or pathology. It is intended that the term encompass any methods for

diagnosis, including, but not limited to assays based on immunoreactivity (e.g., radioimmunoassays, fluorescence immunoassays, enzyme immunoassays), histochemistry, dye retention or binding (e.g., fixing of dyes such as Congo Red), nucleic acid based diagnostic methods (e.g., identification of nucleic acid sequences associated with disease or pathology), etc.

EXPERIMENTAL

The following examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: eq (equivalents); M (Molar); μM (micromolar); N (Normal); mol (moles); mmol (millimoles); µmol (micromoles); nmol (nanomoles); g (grams); mg (milligrams); µg (micrograms); ng (nanograms); l or L (liters); ml (milliliters); µl (microliters); cm (centimeters); mm (millimeters); μm (micrometers); nm (nanometers); °C (degrees Centigrade); U (units), mU (milliunits); min. (minutes); sec. (seconds); % (percent); kb . (kilobase); bp (base pair); PCR (polymerase chain reaction); β-(1-42) amino acids 1-42 of β-amyloid peptide; Tris (tris(hydroxymethyl)-aminomethane); BSA (bovine serum albumin); Fisher (Fisher Scientific, Pittsburgh, PA); Sigma (Sigma Chemical Co., St. Louis, MO.); Promega (Promega, Corp., Madison, WI); Perkin-Elmer (Perkin-Elmer/Applied Biosystems, Foster City, CA); Senetek (Senetek, PLC, Maryland Heights, MO); Boehringer Mannheim (Boehringer Mannheim, Corp., Indianapolis, IN); Stratagene (Stratagene, Inc., La Jolla, CA); and NEB (New England Biolabs, Beverly, MA). Unless otherwise indicated, the restriction enzymes used in these Examples were obtained from NEB. C. elegans is available from the Caenorhabditis Genetics Center, at the University of Minnesota, St. Paul, MN.

5

EXAMPLE 1

Assembly of the β -(1-42) Minigene

30

25

5

10

15

20

In this Example, the β -(1-42) minigene used in subsequent experiments (*See*, Example 3) to produce transgenic *C. elegans* engineered to express amyloidic human proteins was constructed. As described below, these animals contained constructs in which the muscle-specific *unc-54* promoter/enhancer of *C. elegans* drove the expression of the appropriate coding regions derived from human β -amyloid cDNA clones, as described by C.D. Link, "Expression of human β -amyloid peptide in transgenic *Caenorhabditis elegans*," Proc. Natl. Acad. Sci., 92:9368-9372 [1995]).

5

10

15

20

25

30

First, a minigene encoding the 42 amino acid β-amyloid peptide derived from human amyloid precursor protein cDNA was assembled. The artificial signal peptide coding sequence of vector pPD52.81 was amplified under standard conditions, using primers "SP-up" (5'-CGGGATTGGCCAAAGGACCC-3')(SEQ ID NO:1), and "SP-down" (5'-CCGGTACCTGCTGGTGCCAGCAAGAT-3')(SEQ ID NO: 2), cleaved with *NheI* and *KpnI* restriction enzymes, and inserted between the unique *NheI* and *KpnI* sites of vector pPD49.26, to produce the construct "pCL2." This process resulted in a reengineering of the signal peptide, such that the signal peptide cleavage site, as predicted by the consensus of von Heijne (G. von Heijne, Nucl. Acids Res., 14:4683-4690 [1986]), occurred immediately after the Gly-Thr dipeptide encoded by the *KpnI* site.

A 146-bp fragment encoding amino acids 1-42 of β-amyloid, and which contained an artificial stop codon, was amplified using standard methods, from human β-amyloid precursor protein cDNA clone p4T4B (P. Ponte *et al.*, Nature 331:525-527 [1988]), by using primers "β-1-42-up" (5'-GGGGGTACCGATGCAGAATTCCGACATGA-3') (SEQ ID NO:3), and "β-1-42 down" (5'-CCCGAGCTCACGCTATGACAACACCGCCAA-3')(SEQ ID NO:4). The amplification product was cleaved with *KpnI* and *SacI*, and inserted between the unique *KpnI* and *Sac I* sites of pCL2, to generate "pCL3."

The signal peptide/β-(1-42) minigene fragment was removed from this plasmid by digestion with *Nhe*I and *Sac*I, and inserted between the unique *Nhe*I and *Sac*I sites of pPD30.38, to produce "pCL12." The sequence of the β-(1-42) minigene was confirmed by dideoxy DNA sequencing of the coding strand only, by techniques known in the art. The sequence of pCL12 (SEQ ID NO:5), is shown in Figure 2. A graphic map of pCL12 is provided at Figure 3. Figure 4 shows the DNA and amino acid sequences (SEQ ID NOS:

6 and 7, respectively) of pCL12 from nucleotide 1071 through 1253 (i.e., the β (1-42) nucleic acid and amino acid sequences) present in the construct.

5

10

15

20

EXAMPLE 2

Construction of the hsp/GFP Reporter

In this Example, the expression vector ("pCL25") containing the hsp/GFP reporter which drives the expression of the GFP-TT gene was constructed. This vector was used in subsequent experiments (*See*, Example 3) to produce transgenic *C. elegans* was constructed. GFP-TT is a modified form of GFP that contains Ser65Thr and Ile167Thr (the numbering is relative to the sequence of the wild-type GFP). The substitutions present in GFP-TT produce a protein which results in the production of a much brighter fluorescence than the wild-type GFP.

pCL25 was constructed as described below. A 431 bp *HindIII-BamHI* fragment containing the hsp16-2 promoter was excised from the plasmid "pPD49.78," by digestion with *BamHI* and *HindIII*. This 431 bp *HindIII-BamHI* fragment was inserted between the *HindIII* (nucleotide #1) and *BamHI* (nucleotide #31) sites of the GFP-TT gene in the plasmid "pGFP-TT," (available from Yishi Jin, at the University of California, Santa Cruz), using T4 ligase (Promega). The pGFP-TT plasmid contains the coding regions for GFP-TT inserted into a Tu61 backbone. The DNA sequence of pGFP-TT (SEQ ID NO: 8) is shown in Figure 5. A schematic map of pGFP-TT is shown in Figure 6, in which unique restriction sites are shown. The pGFP-TT plasmid was digested with *HindIII* and *BamHI*, and the 431 bp fragment containing the hsp16-2 promoter was inserted, in order to generate the stress-inducible expression vector "pCL25." The DNA sequence of pCL25 (SEQ ID NO:9) and restriction map are shown in Figure 7. Figure 8 provides a schematic map of pCL25 in which unique restriction sites are shown.

5

10

15

20

25

30

EXAMPLE 3

Construction of Transgenic Animals

In this Example, the transgenic parent animals were produced, with one line expressing the β(1-42) minigene (designated as "CL2005") and the other line expressing the hsp/GFP reporter (designated as "CL2070"). For both lines, the transgenes were introduced into *C. elegans* by gonad microinjection as known in the art and described by Mello *et al.* (Mello *et* EMBO J., 10:3959-3970 [1991]). Marker plasmid pRF4 containing the gene (*rol-6*[su1006]) (SEQ ID NO:9)(pRF4 carries a 4 kb *Eco*RI fragment of *C. elegans* genomic DNA containing the *rol-6*[su1006] collagen gene in the Bluescribe vector [Stratagene]; *See*, Mello *et al.*, *supra*; and Kramer *et al.*, Mol. Cell. Biol., 10:2081-2090 [1990]) was coinjected with the constructs, at approximately 100 ng/μl for each plasmid, into morphologically wild-type animals, and Roller transgenic progeny were recovered. The (*rol-6*[su1006]) gene (SEQ ID NO:9), contained within the pRF4 plasmid is a mutated *C. elegans* collagen gene, the expression of which produces the dominant, distinctive "Roller" phenotype. Figures 9 and 10 show the DNA and amino acid sequences of the *rol-6* gene used in this Example, respectively.

 \mathcal{L}

Transmitting lines were established and maintained by selection for the Roller marker phenotype. Transgenic animals produced in this manner maintain the injected DNA as an extrachromosomal, multicopy array of variable mitotic and meiotic stability. Strains containing chromosomally integrated transgenes were recovered by irradiation of lines containing extrachromosomal transgenic arrays with 7000 rad (1 rad = 0.01 Gy) of gamma rays from a Cesium-66 source. Progeny of irradiated animals were then screen for 100% transmittance of the marker transgene. The transgenes in both the CL2005 and CL2070 lines were chromosomally integrated, and were 100% stable.

Transgenic (i.e., as indicated by the Roller phenotype) animals produce both transgenic (i.e., Roller) and non-transgenic (i.e., non-Roller) progeny. These non-transgenic progeny were found to serve as good internal controls for phenotypic and immunohistochemical comparisons. The expression of GFP in the dual transgenic animals can be detected in less than 24 hours after the upshift of animals from 16 to 25.5°C, at all

stages of development, from late embryonic to adult. Expression of GFP resulted in the production of green fluorescence in muscle cells; intense tissue-specific expression of GFP was observed. The fluorescence can be observed using compound or dissecting epifluoroescence microscopy (*i.e.*, with standard fluorescein excitation and emission filters). It is also contemplated that the fluorescence is observable by use of fluorimeters and cell sorters.

EXAMPLE 4

Immunohistochemistry of Transgenic Animals

10

15

5

In this Example, immunohistochemistry was used to confirm the transmittance of the chimeric constructs in large populations of putative integrated lines.

As described by Link (Link, 1995, *supra*), whole mount specimens were prepared by fixing animals in 4% paraformaldehyde and permeabilizing them with 2-mercaptoethanol collagenase as known in the art and described by Link *et al.* (Link *et al.*, Genetics 131:867-881 [1992]). Three antibodies were used (polyclonal rabbit anti-β peptide antibody, Boehringer Mannheim; mouse monoclonal antibody Ab 4G8, available from Senetek; and monoclonal 4.1, a gift from B. Cordell at Scios Nova). The monoclonal 4.1 recognizes residues 8-15 of β-peptide.

20

As was observed by Link (Link, 1995, supra) animals transgenic for the unc- $54/\beta(1-42)$ minigene construct contained muscle-specific deposits of anti- β peptide
immunoreactivity. Non anti- β -immunoreactivity was observed in control animals (*i.e.*,
non-transgenic, wild-type C. elegans tested concurrently). Although C elegans is reported
to contain a homolog of the β -amyloid precursor protein gene (See, Daigle and Li, Proc.
Natl. Acad. Sci., 90:2045-2049 [1993]), this sequence does not contain an apparent β peptide domain. Thus, it would not be expected, nor did it show, cross-reactivity with the
anti- β antibody used in these experiments. For dually transgenic animals, immunoreactive

30

25

Next, to determine whether the immunoreactive deposits observed in the $unc-54/\beta$ -(1-42) strains displayed the tinctural properties of classic insoluble β -amyloid, transgenic

deposits accumulated in the body wall of the animals.

strains were fixed and stained with thioflavin S, a fluorescent amyloid-specific dye, as known in the art, and described by Guntern and Bouras (R. Guntern and C. Bouras, Experientia 48:-10 [1992]). Thioflavin S-reactive deposits were found in all strains containing the $unc-54/\beta-(1-42)$ minigene constructs, but not in control wild-type animals.

5

Whole mounts of fixed dual transgenic animals stained with coumarin-phallacidin (a muscle-specific probe) and anti- β peptide antibody. Muscle-specific β peptide deposits were observed, and showed a qualitative correlation with the level of GFP expression in these muscle cells (*i.e.*, cells that contained more β -peptide deposits were more green than cells with less β -peptide deposits).

10

In addition, Congo Red and Chrysamine G, two dyes known to interact with β -amyloid and have been reported to interfere with its aggregation are tested on the dually transgenic animals. Anti-oxidants (e.g., vitamin E and ascorbate) are also tested. In these experiments, interference with aggregation of β -peptide, or interaction with β -amyloid are observed.

15

20

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

CLAIMS

What is claimed is:

5

- 1. A transparent animal comprising:
 - (a) a toxic transgene;
- (b) an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable *in vivo*.

- 2. The transparent animal of claim 1, wherein the inducible promoter operably linked to the reporter gene is stably integrated into the genome of the animal.
- 3. The transparent animal of claim 2, wherein the inducible promotor is induced by disruption of cellular function.
 - 4. The transparent animal of claim 2, wherein the inducible promoter is a heat shock promoter.
- 20 5. The transparent animal of claim 2, wherein the toxic transgene is β-peptide.
 - 6. The transparent animal of claim 2, wherein said transparent animal is selected from the class Nematoda.
- 7. The transparent animal of Claim 6, wherein said transparent animal is Caenorhabditis elegans.

8. The transparent animal of Claim 2, wherein said genome comprises SEQ ID NO:5 and SEQ ID NO:8.

- 9. The transparent animal of Claim 2, wherein said reporter gene is green fluorescence protein (GFP).
 - 10. The transparent animal of claim 1, further comprising a heterologous gene marker.
 - 11. A method for producing a dually transgenic non-human animal comprising:
 - a) providing:

10

15

20

- i) a first and second non-human animal; and
- ii) a first toxic transgene; and
- iii) a second transgene, comprising a reporter;
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal, and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
- c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic animal, wherein said toxic transgene and said reporter are expressed.
- 12. A dually transgenic non-human animal produced according to the method of claim 11, wherein the toxic transgene is β -peptide.
- 13. A dually transgenic non-human animal produced according to the method of claim 11, wherein said non-human animal is transparent.

14. The dually transgenic non-human animal of claim 13, wherein said non-human animal is a nematode.

- 5
- 15. The dually transgenic non-human animal of claim 14, wherein said nematode is *Caenorhabditis elegans*
 - 16. The method of claim 11, wherein said reporter is green fluorescent protein.
- 10
- 17. The dually transgenic non-human animal produced according to the method of claim 11, wherein said first transgene comprises pCL25.
- 18. The dually transgenic non-human animal produced according to the method of claim 11, wherein said second transgene comprises pCL12.

15

- 19. The dually transgenic non-human animal produced according to the method of claim 11, further comprising a heterologous marker gene.
 - 20. A method for testing compounds for toxicity, comprising:

20

- a) providing:
- i) a dually transgenic non-human animal expressing a toxic transgene and reporter gene operably linked to a promoter inducible by the toxic transgene;

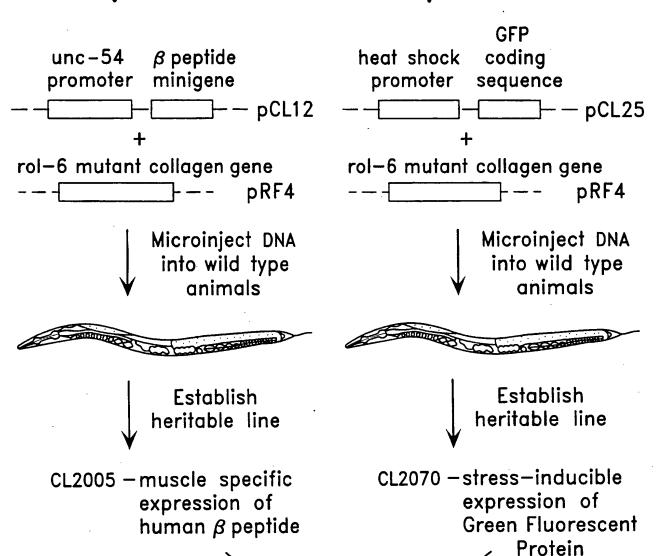
- ii) a composition comprising a test compound in a form suitable for administration such that said compound is bioavailable in the cells of said nonhuman animal; and
 - b) administering said test compound to said non-human animal.

21. The method according to claim 20, wherein the toxic transgene is β -peptide and the reporter gene is GFP.

- 22. The method of Claim 21, further comprising c) measuring a reduction in the fluorescence of said non-human animal and thereby identifying a compound as therapeutic.
- The method of Claim 20, wherein said compounds inactivate said β-peptide
 expressed by said dually transgenic animal.

DNA injection mixture 1:

DNA injection mixture 2:



Mate lines together, recover dual transgenic animals by classic genetic techniques

Fig. 1

Fig.	2Δ
LIE.	CA

Fig. 2B

Fig. 2C

Fig. 2D

Fig. 2E

Fig. 2F

Fig. 2G

Fig. 2H

Fig. 2I

Fig. 2J

Fig. 2K

Fig. 2L

Fig. 2M

Fig. 2N

Fig. 20

Fig. 2P

Fig. 2Q

Fig. 2R

Fig. 2S

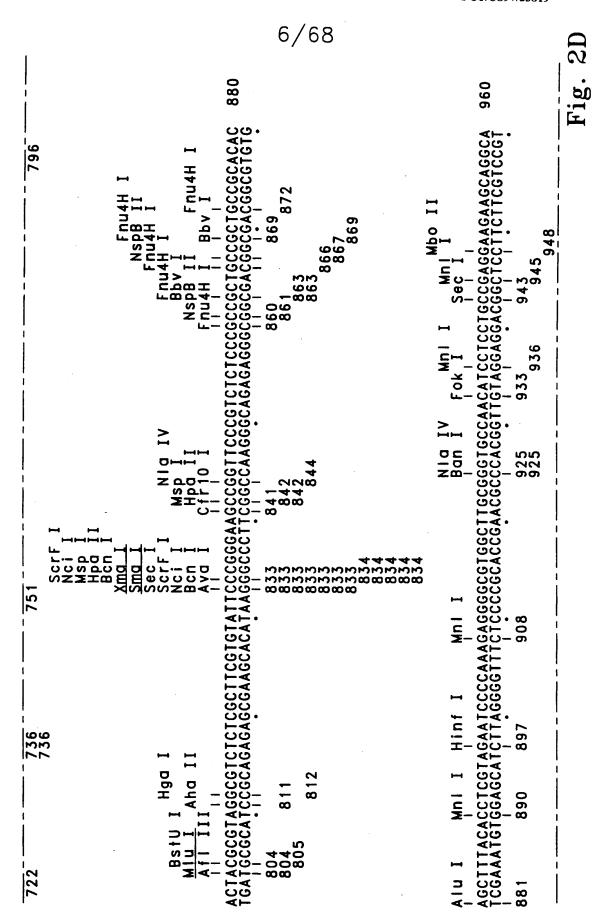
Fig. 2T

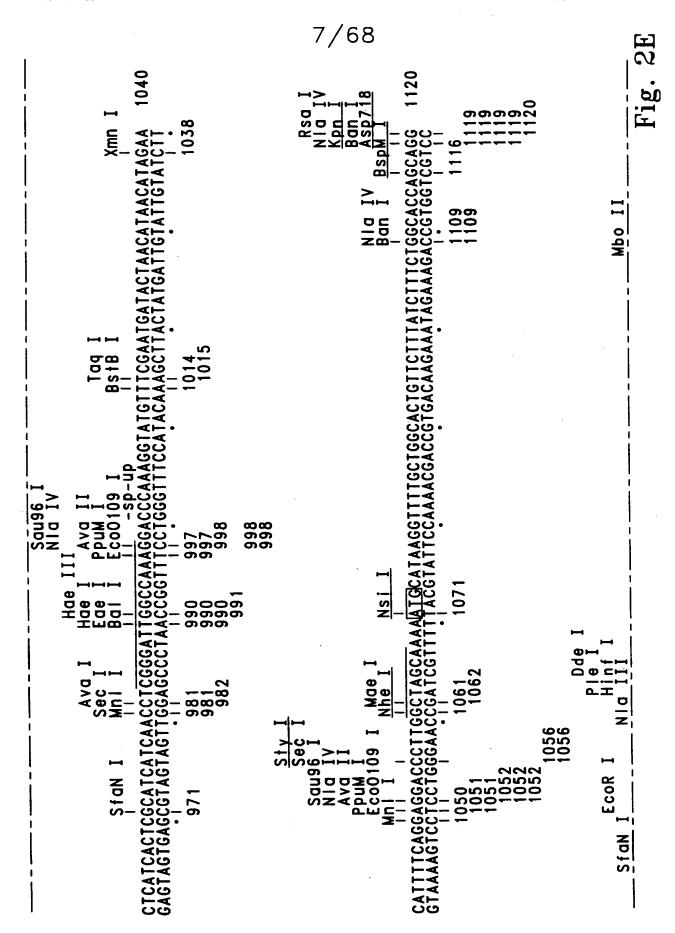
Fig. 2

	3/	⁷ 68		2A
I I 80	160	240	11 II Hinf I - 320	五 五 5
Alu I Mbo II Mbo	Ž5 ATCCCATTCTCTCATCAATTGAGTGGG TAGGGTAAGAGAGTAGTTAACTCACCC	BSP1286 I BSP1286 I BSP1286 I CTCTGAACCATCTTACATTACACTGTGGATGCCCCACAGGCTCCCTTGCATCAGATACTGCCATTGGGGATGGCA GAGACTTGGTAGAATGTAATGT	Squ3A I Mbo II BstY I BstY I Bgl II I Mde II Mbo II Mbo II AGAAAAAGAAAAGAAAGAAGATCTGAAGAAGAAGAAGAAAGA	TCTTCTCTCTTCCALAACACICCTALAAAAAAATTTTTTTTTT

I		4/68		_2B
320	∢ ⊢ (MSE I PQ I inc II TTAA 480 AATT 76 76 477	I Mbo II GCTG 560 CGAC 560	Fig.
308 308 310 310 311 311	MS CATAAAATT GTATTTTAA 39	AAGGGCAAG TTCCCGTTC	Mn I Mn I I CAAGAGAG GTTCTCCTC 551	 - - - - - - - -
299 305	Bsm I GAATGCAACTGAA CTTACGTTGACTT 375	AAAGAGTAATGCA TTTCTCATTACGT	GCTCCTAATA	
-2	O II I Mse I E AGAACTTAAAGGG TCTTGAATTTCCC 368	GGTTCTCTGTAA/ CCAAGAGACATT	MGG I 3A I I I I HPh I CTAGTGGAAC GATCACCACTTGG	
275	MDG BS+B BS+B BS+B ATATATTCGA/ ATATATTCGA/ 358 358 359 364	GGTTTTCACT	Sau Mbo Dpn BstY Mbo II 1 ATCTGAAGAT TAGACTTCTA	•1
- 19	De III TAACAGGGGTC ATTGTCCCCAG	GAGTATTTC CTCATAAAAG	Mse I - ATTTGCTTAAA TAAACGAATTT)
- 2	M TGCACAAGTTAGAAG ACGTGTTCAATCTTC 3	Hinf I Hinf I ATTGAATCCTGCAGT TAACTTAGGACGTCA 410 415	TCGTAAATGTATTGA AGCATTTACATAACT	
243	ATTCATI	AAAGGGA TTTCCC1	Dde I CTTAGG GAATCCA	

1	5,	/68		2C
	640	720	800	Fig.
519 519 520 520 520 520 520	CACTGGTTGTATGTAAACATTC GTGACCAACATACATTTGTAAG	9 ∀ ⊢	F II Hgia I Bsp1286 I Ecor I Mbo II Hbo II CATTCTTGAGACTTCCACTATCT GAAAAGAAGGAGGTTACCTTAAGAACTCTGAAGGTGATACCTTAAGAACTCTGAAGGTGATAGAAAAAAAA	77





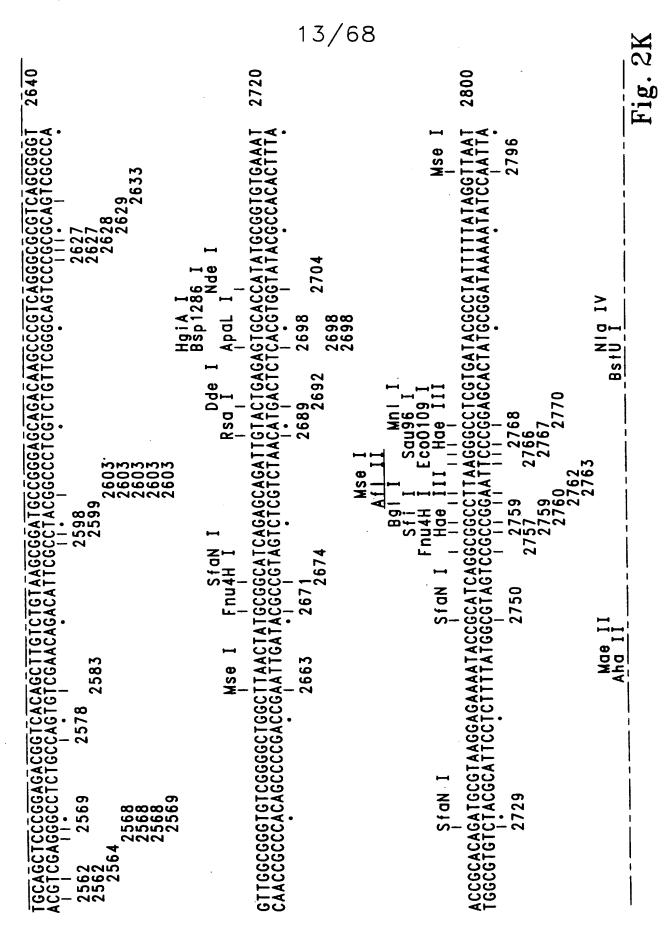
		8/68		, S, F,
	1200	A I I 1280	I 1360	Fig.
	TGTGGGTT ACACCCAA	Sau3/ Mbo Mbo CATCAGAT GTAGTCTA 1278 1278 1278	HgiA I Bsp1286 I CTGCTCC ACACGAGG 1354 1354	-
	TGCAGAAGA ACGTCTTCT 1188	Fnu4H I Hqe III I di II di II dg II CCCCCTGT CCCGCCGCTGT 263 263 263 263 263 263 263 263 263	1266 AAAGAGGG AAAGAGGG	∑
	GGTGTTCTT	u I Sfan A I Sfan 1286 I E 11	N P I I N S P I I N S P I I N S P I I I N S P I I I I I I I I I I I I I I I I I I	n l
	TCAAAAATT AGTTTTTAA	A T G C C C A C T C C C C A C T C C C C A C T C C C C	Fok I TCAACATCC AGTTGTAGG	
	AGTTCATCA TCAAGTAGT	GTGTTGTCA CACCACGT		H-
	AGGATATGA TCCTATACT		Dde CTAA AGATT 1309	0 N-
•	ACATGACTC TGTACTGAG 1140 1143	ZH - ŬŬ	Mn! I CCTCTGACT GGAGACTGA 1299	
	Peptide GAATTCCG CTTAAGGC 1131	TGCAATCA	HinP I Hha I s tu I GCGCCCTG GCGGGCAC 1290	1
	Start of B TACCGATGCA ATGGCTACGT	Tth111 II CAAACAAAGG GTTTGTTTCC 1201	CGCCATCTCC B8 GCGGTGGAGG	

		9/	ر 168	ֻ כ
1440	1520	1600	11 1680	
CACCCCTATTTTGTTATCAAAAACTTCTTCTTAATTCTTTGTTTTTTAGCTTCTTTTAAGAAAAAACGAAAAAAAA	MSe I Ase I Actitacccatcaaaaatacacattaattaattaattaatt	ATAATTCTATCCCAAAATCTACACAATGTTCTGTGTACACTTCTTATGTTTTTTTT	ATCATAGAAAAACCGCACCAAAATACCTTATCATATGTTACCATTTCAGTTTATGACCGCGAATTTTTATTTTCTTCGCATATCTTCGCAATACCAGTCAAATACCGCGAATTTTTTTT	

	10/68		
1840	1920	2000	Fig.
	Sfan I Sausa I Hgia I hho I Hgia I hho I Hgia I hho I Hgia	se I Xmn I TAATGACAGAATACATTCCCA ATTTACTGTCTTATGTAAGGGT 975 1985 Mse I TTAATTTGCAAAATTACTTA AATTTAAACGTTTTTAATGAAT	004 2055 2017 2017 2075 Hinp I Bstu I

		11/68		21
2160	2240	2320	2400	Fig.
FCOR I HHG I MGE III THITTCIGCATTATIGIGITTCCGGCTATATTAGTGTTTGTGTTTGTGTTTTCCGGCTATATTAGTGTTTTGTTTG	Hinf I BS+B I BS+B I BS+B I Hinf I A HINF I BS+B I HINF I A HINF I	- HA.	Mae III Aha I Aha I HinP I Hae III Aqi I Aqi I Hha I Hae III Hae III AATGICGAAATGACGTCATCTGCGCGGAATCATATTCGCCTTTTTTTT	

	12/68	·	—
၂ ဥာ ပုဇ္ပ	Mn I I NSPH I I NSPH I I NSPH I NSPH I NSPH I NSPT	NspB II Hga I Bs†U I HinP I Hha I	Fig. 2J
49	aug6 I ug6 I ug6 I I I I b 1286 I HinP I h 1 I B H	Scrf I Noi I Noi I Hebri I Fok I	
I I TATAGTTATA	Spe I	MSP I ScrF I ScrF I ScrF I Bcn I Alu I Bby I Hpg II Mge III	



		14/68			16
2880	2960	3040 A T	1 I I 3120		ر ع ا
TCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTG1 AGTCCACCGTGAAAGCCCCTTTACACGCGCCTTGGGGATAAACA 2854 2854 2855 5	TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA ACTCTGTTATTGGGACTATTTACGAAGTTATATAACTTTTCCTTCT ACTCTGTTATTGGGACTATTTACGAAGTTATAACTTTTCCTTCT ACTCTGTTATTGGGAGAGAAGTAAACTTTTCCTTCT 2940 2953	TAAAACGGAAGGACAAAACGAGTGGGTCTTTGCTCACCCAGAAACGAAAAAAAA	I MSPB II AIW I NSPB II AIW TCAACAGCGGTAAGAT AGTTGTCGCCATTCTA	043 3056 3065 3075 3087 3100 3109 31 3062 3075 3093 3100 3109 311 3065 3075 3093 3101 31	Mae II

	15/68		2M
3200	3280	3360	Fig.
Hhd I HAG I HAG I HAG I TACTACTCGTGAAAATTTCAAGACGATACACCGCGCCATAATAGGG 3160 3168 3186 3187	I Scall Main Scall Scall Main Main Main Main Main Main Main Main	Fnu4H I Ede II Enu4H II Enu4H II Enu4H II Enu4H II Enu4H II Enu4H II EDE II Enu4H II EDE II E	
CCTTGAGAGTTTTCGCCCCGAAGAAGGTTTTCCA GGAACTCTCAAAAGCGGGGCTTCTTGCAAAAGGT 3140	Scrf I Nci I Msp I Hsg I And I I And I I CATATGACGCCGGCCAGCCAGCCGC CATAACTGCGCCCGTTCTCGTTGCCGCGC CATAACTGCGCCCGTTCTCGTTGAGCCAGCGCGCGCGCGC	T-000 8-00-8	Squ3A_1

	16/68		Z
- Squ3A I Mbo I Dpn I - 3440 3440 3440	Mae II ACG 3520 TGC 3518	AGT 3600	
GCACAACATGGGGGATCATGTAACTCGCCCGTGTTGTACCCCCTAGTACATTGAGGGGATCATTTGAGGGGGGGG	ACACCACGATGCCTGTAGCAATGGCAACA	Mse I Fok I CATTAGACTGGATGGAGCGGATAAGGTTAATTATCTGACCTACCT	Hph I Msp I — — — — — — — — — — — — — — — — — —
	III Alu I GAGCTGAATGAAGCCATACCAAACGACGAGGTG CTCGACTTACTTCGTATGGTTTGCTGCTCGCAC CTCGACTTACTTCGGTATGGTTTGCTGCTCGCAC 3456 3456 3456	HinP I Hhd I Hhd I Kei I Nci I Nci I Nci I Nci I Nci I Nci I Scrf I Scr	<u>ရ</u>

	17/6	88	20
3680 3680	3760	3840	다. 19.
6661C1C CCCAGAG 366	GCAACT CGTTGA	TTACTC AATGAG	II
A C C C A	Hinf I GAGTCAGC CTCAGTCAG 3748	CCAAGT.	
HP	ACGGG TGCCC	I GTCAGA CAGTCT	L
AGACCTO 3 659 3 3 3 5 9	CTACACG GATGTGC	Mae II FTGGTAACT AACCATTGA 3817	Sau3A Mbo I Dpn I Bs+*
GATAAA	AGTTATO •	I AGCATTO TCGTAAC	A I I I I Mb
ATTGCT	TATCGT ATAGCA	MSe TGATTA ACTAAT 380	Sau 3/
TGGTTT SACCAAA	Mn! I CCTCCCG GGAGGGC/ 3717	Mn I I I I I I I I I I I I I I I I I I I	Mse I
GGCTGGC CCGACCG 29	Mn1 GTAAGCCCT(CATTCGGGA(3717	N D Ban ATAGGTG TATCCAC 3794 3794	1
11 HP 1 HP 11CC AAGG AAGG 36	I GATG CTAC	Dde I I CTGAG GACTC	3786
'a 6 00 6 66 6	ACCCC 3700 3700 3700	Sau3 Mbo Dpn CAGATC GTCTAG 3781 3781	X Se I
TCTGCG AGACGC 361	nu4H I Ibv I CAGCAC GTCGTG	AATAGA	
1096 I 0 II 16616A 17616A	TCATTG AGTAAC AGTAAC 33	I GAACGA CCTTGCT	
20 A A C G C C C C C C C C C C C C C C C C	GCGGTA	Fok ATGGAT TACCTA 3763	

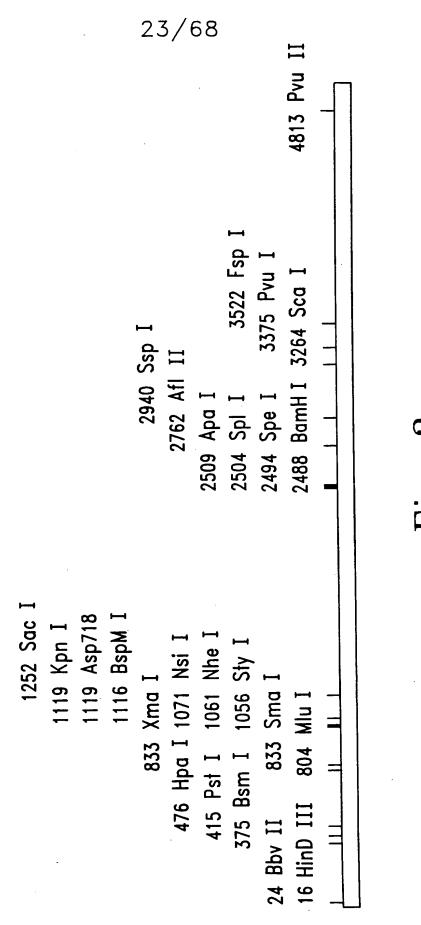
The first indicate The fin
Mac Mac Mac Mac
MSe Hph Hph Hph
I
TATATACTITAGATICATITA ATATATGAATCTAACTAAAT 3859 3859 3859 WAS II WAS II WAS II CAAAATCCCTTAACGTGAGTT GTTTTAGGGAATTGCACTCAA 3931 3931 HINP I FRU HINP I BBY HINP I BBY HINP I BBY HAAAAGGGGGGATTAGACGAA

		19/68		29
	യ ത		4320	Fig.
4067 4067 4070 4070	Mae I HACTGTCCTTCTAGTGTAGCC ATGACAGGAAGATCACATCGG	Fnu4H I Fnu4H I Fnu4H I A I WN I Bbv I H	NSPB II Fnu4H I Bbv I HinP I Hha I CGCAGCGTCGGCTG CCGTCGCCGAC CGCTCGCCAGCCGAC 	295
	HinP I Hha I CAGCAGAGGCGCAGATACCAAAI GTCGTCTCGCGTCTATGGTTTA	ni I CTCGCTCTGCTAA GAGCGAGACGATT 198	MSP I HPG II HPG II HPG II HPG II HPG II HPG II I HPG II HPG II I	
4020 4022	MDE III CCGAAGGTAACTGGCTT GGCTTCCATTGACCGAA	GAACTCTGTAGCACCGC CTTGAGACATCGTGGCG	Scrf I Nci I Msp I Hpd II Ple I Bcn I Hinf I TTACCGGGTTGGACTCA AATGCCCAACCTGAGT 4257 4265 4257 4265	25
4009 4010	ACCAACTCTTTTT(TGGTTGAGAAAAA	GCCACCACTTCAA	GATAAGTCGTGTC CTATTCAGCACAG	

	20/68		2R
4400	4480	4560	I I Fig.
၂ ၁၅၅ -	CAGGG GTCCC 4755 6755 6755 6755 6755 6755 6755 6755	Scrf I Ecor II Bstn I GGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG CCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGTGGAGCTGAACTCGCAGCTAAAAAAAA	Hae III Scrf I III Scrf I III Scrf I III Scri I I Scrf I III Scri I I I I I I I I I I I I I I I I I I I

	21/68	SS SS
4640	4720 II 4800	Fig.
GGGCGGAGCCTATGGAAAACGCCAGCAACGCGCCTTTTACGGTTCCTGGCCTTTTGCTGGCAACGACAAACGACGGAGGACGGAGGAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGACGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAACGAGGAAAAAA	- PA 00	Mse I Alu I Ase I NspB II

```
22/68
4880
      4855
4855
                                MSP
Hpd
                                                       4976
                                          89844
88997
778
                  4891
4891
                          4806
4807
         4801
```



F1g.

24/68

T CTG GCA CCA GGT ACC GAT GCA

T CTG GCA CCA GGT ACC GAT GCA

1 31

T CAA AAA TTG GTG TTC TTT GCA GAA

S gin lys leu val phe phe ala glu

51

C ATG GTG GGC GGT GTT GTC ATA GCG

tu met val gly gly val val ile ala 1101 /
ATC TTT C
ile phe 1
1161 /
CAT CAT C
his his ç
1221 /
GGA CTC /
gly leu r 11 AAG GTT TTG CTG GCA CTG TTC TTT ATIS Iys val leu leu ala leu phe phe il 11 CCA CAT GAC TCA GGA TAT GAA GTT CA GGA TAT GAA GTT CA A1 CAT AAC AAA GGT GCA ATC ATT GG al gly ser asn lys gly ala ile ile gl

met 1071 /
ATG CAT /
met his 1131 /
GAA TTC glu phe 1191 /
GAT GTG asp val 1251 /
TGA OPA

25/68

Fig. 5A	
Fig. 5B	
Fig. 5C	·
Fig. 5D	
Fig. 5E	
Fig. 5F	
Fig. 5G	
Fig. 5H	
Fig. 5I	Fig
Fig. 5J	angle Fig.
Fig. 5K	
Fig. 5L	
Fig. 5M	
Fig. 5N	
Fig. 50	·
Fig. 5P	
Fig. 5Q	
Fig. 5R	

```
26/68
```

```
27/68
                                                                                                                   Fig.
                                                              160
                                           RSG
NIG
KPD
Ban
ASP 7
                                                                          ထထထထ
                                                                                                               86
                                                                                                 Hph
32
                                                                        00-
00-
00-
00-
00-
00-
00-
00-
                                                          LEX.
```

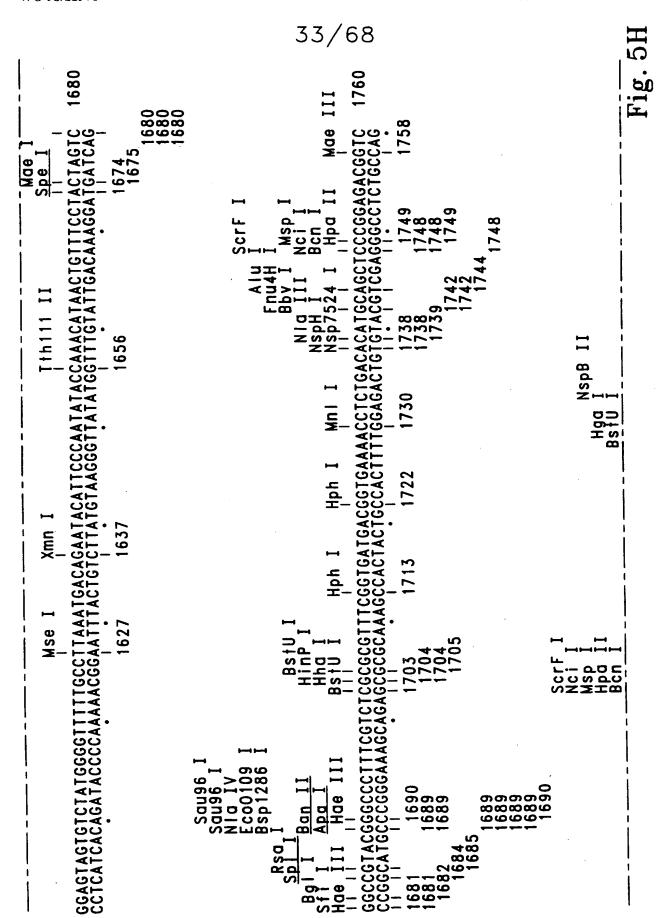
	2,8	3/68	5C
	320	400	1 480 Fig.
Hae III Eqe I Eqe I Eqe I Equipment Hair Equi	- CC	Nde I Sau3A I Mbo I Mbo I Dpn I CTITCGCITTECAATGCTTTTCAAGATACCAGATCATATGAAACGGCATGAAAGTTCTCAAGAGTGCCATGCCC GAAAG+GAAAAGTTCTATGGGTCTAAGAAAGTAACGGGTACTGAAAAAGTTCTCACGGTACGGG 359 359 359 359	GGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGG CCCTTGATGTTCTGTGCACGTTCAGTTTGAAGG

11 0 30/203 / 1				
	29/	⁷ 68		2D
260	hr167 640	720	800	Fig.
1 A P	∀⊢	AAATACTCCAATTGG TTTTATGAGGTTAACC 3A I	DPN I Alw I Bsty I BstB I AGACCATTACCTGTCCACACAATCTGCCTTTCGAAAGAACGAGAGAGA	

1	30/68		5E
 	880	096	Fig.
784 787 787 760 787 787 787 761 781 781 781 781 781 781 781 781 781 78		09 ACCAACTTGTCTGGTGTCAA TGGTTGAACAGACCACAGTT	

	3	1/68		5F
1040	1120 1120 20	1200	1280	Fig.
Mbo II NSPH I Hgiā I BSP1286 I BSP1286 I SP7524 I BSP1286 I SP7524 I BSP1286 I SP7524 I SP892 I	MSE I MGE III MGE III I MSE I MN! I TAATTICTTTGTTTTTTTTTTAGTCACCTCTAACAATGAATTGTGTGGAGATTCAAAATAGAA 1 AATTAAAGAAACAAAAAAAATTCAGTGGAGATTGTTACTTTTAACACAAAAATAGAA 1 AATTAAAAGAAACAAAAAAATTCAGTGGAGATTGTTACTTTTAACACATCTAAGTTTTTATCTT AATTAAAAGAAACAAAAAAAATTCAGTGGAGATTGTTACTTTTAACACACATCTAAGTTTTTATCTT AATTAAAAGAAACAAAAAAAAAA	MSe I ASe I III CATTAATAATTCTATCCCAAAATCTACACAATG GTAATTATTATTAAGATAGGGTTTTAGATGTTAC II66 1166	AAAAAAACTTTGTAGTATTTTTTTGGCGCACACAAATA AAAAAAACTTTGTAGTATCTTTTTTGGCGTGTGTTTTAT	Mae III Mae II Mae II Mae II Mae II Mae II Nio II Sau96 I And II Nio
Dde I CTAAGTCCA GATTCAGGT 961	Mbo I AACTTCTTC TTGAAGAAG	MSG I TTAATTCG AATTAAGCA	TTCTGTGT/ AAGACACA	e N

				2/6	8						3 C
 GTTAAAATAAAGAAGCGTGCAGACCCGGAGTACTGCAGTTTAGT 	1345 1346 MSe I ASe I	FTGGAATTTTCAATCAAGTGAAGTTTATGAAATTAATTTCCTGCTTT 1440 AACCTTAAAAAGTTAGTTCACTTTCAAATACTTTAAAAGGACGAAA 1424	I jag I Bsp1286 I	AAGAGTTÍCGAGGACGGCGTTTTTCTTGCTAAAATCACAAGTATTGATG TTCTCAAAGGTCCTGCCGCAAAAGAACGATTTTAGTGTTCATAACTAG	79 1520 1481 1520	Dde I	TGGTTTGAGG ACCCAAAÇICC	59 1563 1573		Ede I	Fig. 5G
 ĞĞAAĞTĞAAATAĞTĞĞĞĞT 295		GGA	II 111411	GTTTGTCAAGA CAAACAGTTCT	1464	Sau3A Mbo I Mbo I I ngh I	CAAGAAAGATCGGAAGAAGGTTTGGG STTCTTTÇTAGCÇTTCTTCCAAACCC	526 1537 1537 1537)		



	34/68		51
1840	1920	de I 2000 000	 - - - - - -
ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAAGCCCGTCAGGGCGCTCAGCGGGTGTTGGCGGGTGTCGGGGCTG ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAAGCCCGTCAGGGCGCTCAGCGGGTGTTGGCGGGTGTCGGGGCTG TGTCGAACAGACATTCGCCTACGGCCCTCGTTCGGCCAGTCGCCCACAACCGCCCACAGCCCCGAC TGTCGAACAGACATTCGCCTACGGCCACAAGCCCCGACAGCCCCGAC TGTCGAACAGACAACAGATTGGCCACAGCCCCGACAGCCCCGAC TGTCGAACAGACATTGGCCACAGCCCCTCGTTCGGCCCCACAACCGCCCCACAACCGCCCCGAC TGTCGAACAGACAACAGACAACACAAACACAAACACAAACAA	N	II Squy6 I Eco0109 I Hqe III AGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATATTTTTC AGGCCTCGTGATACGCCTATTTTTTTTTTCCAATTACAGTACTATTATTACCAAAG ICCCGCAGCACTATGCGGATAAATATCCAATTACAGTACTATTATTACCAAAG I 1948 1948 1948 1946 1947	

	35/	68		
2080	2160	2240	0 II 2320	2320 Fio
Mge II Aha II Hinp I Hinp I Aqt III Hha I Hha I ATCTGCACTCACCTTTTCGGGAAATGTGCGCGCTTTGTTTTTTTT	CTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTGAAAAAGGAAGATATGAAAAAGGAAGAGTATTCATAACTTTTTCCTTCATACTCATAAGTTGTAAAAAAAA	CCTTATTCCCTTTTTTGCCTTCCTGTTTTTGCTCACCAGAACGCTGGTGAAAGTAAAGATGC GGAATAAGGGAAAAAAAAAA	PO I BSP1286 I TOO I BSTY I NSPBII BS II Apol I MOO III AIW I NSPBII A ATCAGTIGGTGCACGAGTGGGTTACATGACTGGAACTGGATCTCAACAGGGGTAAGA TAGTCAACCCACGAGTGGCTTACATGTAGAACTGAACAGGTTGTCGCCATTGT	242 2245 2245 2245 2245 2245 2245 2245

	3	6/68		7
1	2400	2480	1 2560	Fio
Scr F I KSp I I Bpd I I I I I I I I I I I I I I I I I I I	77 - 72 - 74 - 75 - 75 - 75 - 75 - 75 - 75 - 75	Sfan I Fok I AGCATCTTACGGATGG TCGTAGAATGCCTACC 1 2466 2475	Sau3A Mbo I Dpn I Pvu I I I CTTCTGACACGATCG GAAGACTGTTGCTAGC 2555 2556 2556	
	GCGCTATTATCCCGTATCCCCTATAGCCCATAGCCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCCCATAGCCCCATAGCCCCATAGCCCCATAGCCCCATAGCCCATAGCCCATAGCCCATAGCCCATAGCCCATAGCCCCATAGCCCATAGCCCATAGCCCATAGCCCATAGCCCATAGCCCATAGCATAGCATAGCATAGCCATAGCAT	. OO	2449 Hae III Gdi II Ede I Fnu4H I 111 ACACTGCGCCAACTTA TGTGACGCCGGTTGAAT 2533 2534	
	- UO- MMN	RSG SCG J CACT GALTGALT CACT GACCAACT CAT CAT CAT CAT CAT CAT CAT CAT	NIG III TAACCATGAGTGATA/ ATTGGTACTCACTAT1 2517	
O I C	CGTGAAATTT CGTGAAATTT CGTGAAATTT 40 2348	I TACACTATTCTCAGAAI ATGTGATAAGAGTCTT/ 2426	Fnu4H I Bbv I TTATGCAGTGCTGCCA AATACGTCACGACGGTA 2506	
II am	TTTTCCAAT AAAAGGTTA 5	Fnu4H CAACTCGGTCGCCGCA GTTGAGCCAGCGGCGT	NIG III CATGACAGTAAGAGAA GTACTGTCATTCTCTT	

	37/68		2T
2640	2720	2800	Fig.
Squ36 I Squ96 I Avg II	CA TO	MSP I Hpg II Scrf I Nci I Mae I Bcn I Ase I Fok I Hpg II Hnd I Avg II Hhg I Av	

		38/68		
I NI DIN	CTG 2880 GAC 2880	ACA 2960 TGT		3040
Fnu4H Bbv I	TCATT CCAGCA AGTAACGTCGT 2872 2872	I Gaacgaaatag Ctt G Ctttatc	CTTTAGATTG GAAATCTAAC	
Bstu I	GGTCTCGCGGTA' CCAGAGCGCCATA' 2860	Fok CAACTATGGAT GTTGATACCTA 2943	TACTCATATATA VATGAGTATATAT	
<u> </u>	CGGTGAGCGTG GCCACTCGCAC 1 42 843 843 845 Ple I	Hinf I CTACACGACGGGAGTCAGG GATGTGCTGCCCTCAGTCC 2928 2928	MGG III GGTAACTGTCAGACCAAGTT CCATTGACAGTCTGGTTCAA 2997	N 20
742 742 742	ATACGACTATTT ATAACGACTATTT	TATCGTAGTTAT ATAGCATCAATA	MSe I CTGATTAAGCATT GACTAATTCGTAA(ž m
P I 277	GGCTGGCTGGTTT CCGACCGACCAAA	1 - I AGG 97	Mn I n I c C C C A A C G C A C A	2974 2978
MS HP HP Guge I I HP I I I I I I I I I I I I I I I I	GCCCTTCC GGGAAGG CGGAAGG 802 802 2803 2803 2803	qu96 I GGCCAGATG CCGGTCTAC B81 2882 Dde I	u3A I o I n I TCGCTGAG AGCGACTC	2961

	39/68	5N
I AG 3120 TC	Fnu4H I Bbv I 1G 3200 3200 3200	Mae III 66 3280 Fig.
Mge I Mse I HTCCTTAACGTG TAGGGAATTGCAC 3111	Bs+U I HinP I Hha I 	CICITITICGGAA
NIG III BSPH I ICTCATGACCAAA AGAGTACTGGTTT 3095	Squ3A I Mbo I Dpn I Alw I Alw I CCTTTTTT CTCTTTTTT 3174 3175 3175 3175 3175	I I Alu I TCAAGAGCTACCAACTCTTTTTCCGAAGG 3
CCTTTTGATAA GGAAAACTATT	88 8 Sau3A I Mbo I I Mbo II BstY I A W I A GGATCTTT TCCTAGAGGA 3163 3163 3164 3164 Sau	MSP I HPG II THH11 II
I BS 1 I HPh I I I I I I I I I I I I I I I I I I I	3069 307 307 307 307 307 307 307 307 307 315 3156 3156 3156	NSPB II ACCAGCGGTGGT
MSe I Dpn Drd I BSTY MSe I Dpn Dpn Drd I BSTY MS I Alw ALTTAAAGGAT AAATTTCCTA 3058 3065 3065 3059	Hga I Hga I GAGCGTCAGACCCC CTCGCAGTCTGGGG	THHIII II CTGCTTGCAAACAAAACCACCGCTACCAGCGGTGG
TAAAACTTCATTI	TTTTCGTTCCACT AAAAGCAAGGTGA 31	T+h111 II CTGCTTGCAACA

							40/	6	8						20
		.80			3360	-			3440					3520	Fig.
	CTTCC	32			AACTCT TTGAGA	7			ACCGG TGGCC	333 444 3333	43		I n	CTTGG	
	AAAAGG				TCAAG AGTTC				TAAGTCGTGTCTTA				A-	ACAGCCAGCT TGTCGGGTCGAA	
	TTGAGA			111	ACCACT TGGTGA				AAGTCG TTCAGC			H 0	^	CACACA	
	:GATGG:	3257		Hde I	AGGCC TCCGG	3339 3340			SGCGAT/			H 0	Apal	AC A	
	GTTCTC		_000		CATCAA		H H		GCTGCCAGTGGCGA CGACGGTCACCGCT		_			700000	
	CGCCACCAAACAAACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCC	•	ა ഗოകക 4000 ഉദ്യവസ		GTAGCCG CATCGGC/		Fnu4H Fnu4H I I Bbv I	٦_ ^q	CCGACGAC	3406 3409 3409	3409			GCTGAACGGGGGG	
	ACAAAC	3241		Mde I	CTTCTAGT	3322	Z ≥		AG.		'	-		SecTesson SecAssons	1
	CACCAA				TACTGTCCT ATGACAGGA			Mae I	⋖⊢	3396		NspB nu4H i I	_ 	AGGGGAGGG TCGGGTGGCG	
	SGTCCC	3230			_ ~				TAATCC	•		L-60 :	 	AAGGCG TTCCGC	
	GCGAT		•		AGATACCAAA TCTATGGTTT				TCTGC.			— і		ACCGGATAAG	
	110011			HinP I Hha I	CCCAC	3298 3298		I - UX-	CTCGC	3378			M de	TAGTTAG	
	<u> GĄCGAACĢTĮTGTTTTTTGGTGGCGATGĢT</u>				CTGGCTTCAGCAGAGCGC GACCGAAGTCGTCTCGCG				TACATA	3378				AGACGAT TCTGCTA	
	VCGTTT	3208			SCCTTC				ACCCCC	•			le I inf I	ACT.	
	GACGA	3202			TAACT ATTGA	٠			GTAGCA(CATCGT(ı	o.∓.	GTTGGAC.	

		41/68		5P
	3600	3680	3760	Fig.
3463 3472 3463 3474 3474 3474 3475	GAAAGCGCCACGCTTCCCGAAGGGA(CTTTCCT) 3571 3572 3572 3572 Sec I	Msp I Fnu4H I HinP I BstN I EcoR II BstN I EcoR II BstN I Stoccottccccccccccccccccccccccccccccccc	Hga I Stan I Stan I GGCGCGCGCGCCGCCGCGCGCGCGCGCGGCGGCGGGGGG	

		42/68		50
	3840	1 3920	4000	Fig
	ပို့ဖွဲ့ -	Fnu4H I Bbv I Hinf I GATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGCGCG	у н	

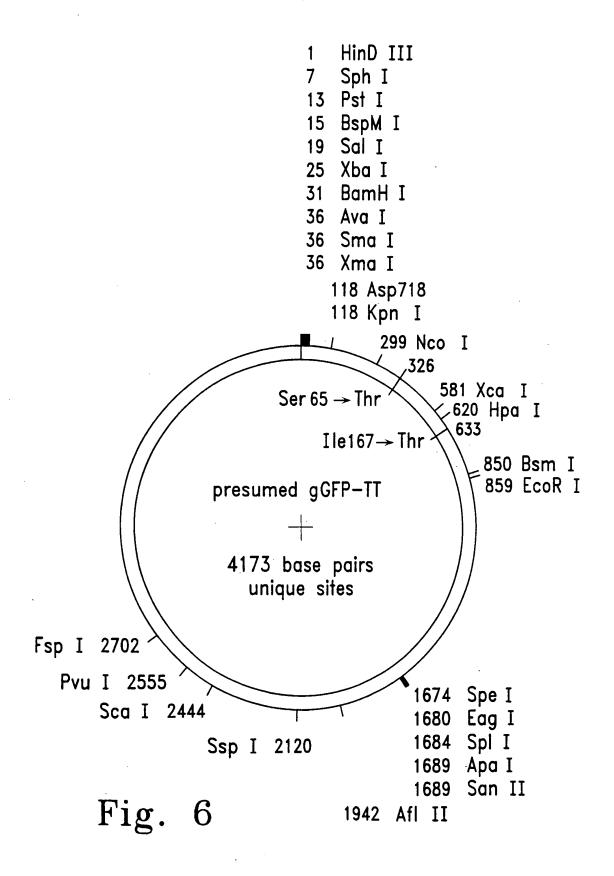


Fig. 7A Fig. 7B Fig. 7C Fig. 7D Fig. 7F Fig. 7F Fig. 7H Fig. 7I Fig. 7J Fig. 7X Fig. 7X Fig. 7X Fig. 7N Fig. 7N Fig. 7N Fig. 7O Fig. 7P Fig. 7P Fig. 7C Fig. 7C	
Fig. 7C Fig. 7D Fig. 7E Fig. 7F Fig. 7G Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7K Fig. 7N Fig. 7N Fig. 7N Fig. 7O Fig. 7P Fig. 7Q Fig. 7R	Fig. 7A
Fig. 7D Fig. 7E Fig. 7F Fig. 7G Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7K Fig. 7N Fig. 7N Fig. 7N Fig. 7O Fig. 7P Fig. 7Q Fig. 7R	Fig. 7B
Fig. 7E Fig. 7F Fig. 7G Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7K Fig. 7N Fig. 7N Fig. 7O Fig. 7O Fig. 7P Fig. 7Q Fig. 7R	Fig. 7C
Fig. 7F Fig. 7G Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7K Fig. 7N Fig. 7N Fig. 7O Fig. 7O Fig. 7P Fig. 7Q Fig. 7R	Fig. 7D
Fig. 7G Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7L Fig. 7M Fig. 7N Fig. 7O Fig. 7O Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7E
Fig. 7H Fig. 7I Fig. 7J Fig. 7K Fig. 7L Fig. 7M Fig. 7N Fig. 7O Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7F
Fig. 7I Fig. 7J Fig. 7K Fig. 7L Fig. 7M Fig. 7N Fig. 7O Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7G
Fig. 7J Fig. 7K Fig. 7L Fig. 7M Fig. 7N Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7H
Fig. 7K Fig. 7L Fig. 7M Fig. 7N Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7I
Fig. 7L Fig. 7M Fig. 7N Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7J
Fig. 7M Fig. 7N Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7K
Fig. 7N Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7L
Fig. 70 Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7M
Fig. 7P Fig. 7Q Fig. 7R Fig. 7S	Fig. 7N
Fig. 7Q Fig. 7R Fig. 7S	Fig. 70
Fig. 7R Fig. 7S	Fig. 7P
Fig. 7S	Fig. 7Q
	Fig. 7R
Fig. 7T	Fig. 7S
	Fig. 7T

Fig. 7

```
46/68
                                                      160
                                     22
22
  SIE
                                                                66
                              തതത
                                                  Sfan
     NSPH I
NSp7524
Sph I
```

1		47/68	ı	7B
 ¦	320	400		Fig.
174 184 209 179 187 209 179 187 210 210	RSG I RSG I PIG I MNI I FRU4H FOUTH I HINT I	MDB I Xba I CAATGTICTAGAAAAAGGTGGAAAATAGTATACCGTTGAAAATAAAT	Nci I Nci I Msp I Hpd II Smg I Smg I Sec I Sec I Nci I Bcn I Avg I Sec I Sec I	

```
48/68
       480
       gtttcgaatgatac
caaagettactatg
⊕ 00 ←
```

·	49/68	02
 taacataacattttcagGAGCACCCTTGGAGGTACGGTAGAAAAAATGAGAAGAACTTTTCAC attgtattgtatcttgtaaaagtcTCCTGGGAACCTCCCATGGCCATCTTTTTACTCATTTCTTTTTACTTCTTTTAAAGAAGAACTTTTCAC attgtattgtatcttgtaaaagtcTCCTGGGAACCTCCCATGGCCATCTTTTTTACTCATTTCTTTTAAAGAAGAACTTTTCAC attgtattgtattgtatcTTTTTAAAGAAAAAGTG 510 523 523 547 547 519 519 510 510 522	HPh I MSe GGTGATGTTAA GCACTACAATT CCACTACAATT 592 599 1ATTTGCACTA ATAAACGTGAT	700 703 703 704 Fig.

	50/68		7E
800	880	096	1040 Fig.
N D III CATGCC GTACGG 195	116AAG AACTTC 880	TTAACCTT TTAACCTT Thr167 Mae I	ATCTGT 034
6A6T6 C1CAC6	TCAAGT AGTTCA	CACAAA 616111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	CAAABC GTTTTG
TTTCAA AAAGTT	CTGAAG	CCTTGGA CGAACCT	17AACTT ATTGAA 121 022
a III ACTGAA '5 Mae I	ACACGTG 1616CAC 1856 858 858	TTGTAAGG	CAAA.
AAC9GCA 1166CGT	ACAAG TGTTC II	CGATGG CATGGG ACCGGGGGGGGGGGGGGGGGGGGGGG	1013
de I – I – I – I – I – I – I – I – I – I	GGGAAC CCCTTG CCCTTG	1 A A A G A A T T T C T C T C T C T C T C T C T C	TTTCT
Sau3 Mbo Mbo Dpn CAGATC GTCTAG 760 760 760	GATGAC CTACTG	⊢∢	999
GATACC	TTCAAA	ACCATA TCCATA TCCATA	ATGCCAC TACCGTC 90
TTTTCAA	TATATIT TATAAA MSe I	AGTTAAA TCAATTT 907	ATGTAGT ATGTAGT 99
TCAATGCT AGTTACGA	ტ Т. ⊏		AA1G1A 11ACA1 982 982
GGTGTTC	CAGGA GTCCT Se I	1178 148 198 198 198 198 198 198 198 198 198 19	AGTGTG
1 1 2 3	A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1	666AAC	ATATTG ATATTG
ACT T T S A A A A A A A A A A A A A A A A	CGAAGG	GTGATA	ATGTTG

	51/68		7F
1120	1200	1280	Fig.
Mae I Mae I Hae IIII Hae IIIII Mae I Hae IIII Hacattere Adam Adam IIII Hae IIIII Hae IIIIIII Hae IIIII Hae IIIII Hae IIIII Hae IIIII Hae IIIIII Hae IIIIIIII Hae IIIII Hae IIIIIII Hae IIIII Hae IIIII Hae IIIIIIIIII	Sau3A I Mbo I Dpn I By	Fnu4H I Bbv I Alu I Nig III NSPB II NSPB II NSPB II NIG III NI	

1

52/68	76
1360 1440 1520	Fig.
1208	

TZ	Fig	
	1920	AGAGTT
		I Mul I I I I I I I I I I
	1840	AIGCICATCGIGAAAAGTITTGGAGTATTTTTGGAATTTTTCAATCAAGTGAAAGTTTATGAAATTAATT
		1747 MSG I ASG I
	760	1726 1739 1750 1 1742 1750 1740 1751
	1760	TTCAGTTTATGACCGCAATTTTTATTTCTTCGCACGTCTGGGCCTCTCATGACGTCAAATC AAGTCAAATACTGGCGTTAAAAATAAAGAAGCGTGCAGACCCGGAGAGTACTGCAGTTTAG
	111	Hae III Mae II Mae II Mni I Aha II Mbo II Sau96 I Aqt II N
		III DIN
	1680	GTICTGTGTACACTICTTATGTTTTTTTTTCTGATAAATTTTTTTTGAAACATCATAGAAAAAACGCGCACACAAAT CAAGACAÇATGTGAAGAATACAAAAAAAAGAAGACTATTTAAAAAAAA
		1558 Rsa I
		521 1540 1552 1552
	1600	
		ASB I BSp1286 I Așe I

		54	/68			71
.	2000	2080		111	2160	ilg.
	AIAACAAAÇAGIICICAAAGCICCIGCCGCAAAAGAAÇGAIIIIAGIGIICAIAACIA 1865 1880 Mbo II A I I Hph I Hph I GGAAGAAGGIIIGGGIIIGAGGICAGGIGAAGGIGAAGGIGAAAGGIICAAACITIÇAAAG	1960 1964 1974 1943 Mae I Mae I Xmn I THIIII Spe I	AAACGGAATTTACTGTTATGTAATGGTTTGTATTGACAAAGGATGATCA 2028 2038 2057 2075 2076	Scrf I Bstu I Hinp I Hha I Hph I Mph I Mni I NspH I Bcn I Stu I Hph I Mni I NspH I Bcn I	CTCGCGCGTTTCGGTGATGAGGTGAAACCTCTGACATGCAGCTCCCGGAGACGGT 21 GAGCGCGCAAAGCCACTTTTGGAGACTGTGTAGGTCGAGGCGTCTGCCA 2104 2114 2123 2131 2139 2150 2159 2105 2105	Fig. 71
	AACGAAAACCCCCAAAGGGGGGGGGGGGGGGGGGGGGG	1938 1938 1938 1938	CCTCATCACAGATACCCCAA Sauge I Sauge I	SY SECTION OF STATE O	GCCGTACGGCCCTTTCGT CGCCATGCCCGGGAAGCA 81 2090	

		55/6	88		7.3
 		2240		2320	Fig.
2140 2149 2149 2143 2150 2145 2149	Hgd I NspB II Bstu I Hinp I Hi	GGCGCGCTGTTGGCGGGTGTCGGGGCT CCGCGCCACAACCGCCCACAGCCCCGA C2208 2208 2208	2214	TATGCGGTGTGAATACCGCACAGATGCGTAAGG ATACGCCACACTTTATGGCGTGTCTACGCATTCC 2310 85	
	2086 2090 2091	TETETETAAGCGGATGCCGGGAGCAGAGCCCGTCAGAAGCCCGTCAGAAGCCCGTCAGAAGCCCGTCAGAAGCCCGTCAGAAGCCCGTCAGAAGCCCGTCAGAAAGCCCGTCAGAAAGCCCGTCAGAAAGCCCGTCAGAAAGCCCGTCAGAAAAAAAA	2184 2184 2184 2184 HgiA I Bsp1286 SfaN I Dde I Nd	AACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCA TTGATACGCCGTAGTCTCGTCTAACATGACTCTCACGTGGT TTGATACGCCGTAGTCTCGTCTAACATGACTCTCACGTGGT 44 2252 2279 22	

	56/68		ļ	ンド
2400	2480	2560	40	F10
TAATGGTTT ATTACCAAA	TCAAATATG AGTTTATAC	ACATT TGTAA	TAAAAGATG ATTTTCTAC 2637	
BSPH I I I I I I I I I I I I I I I I I I I	TTCTAAATACAT AAGATTTATGTA	TATGAGTATTCA ATACTCATAAGT HDh T	CGCTGGTGAAAG GCGACCTTTC	
MSB CCTATTTTTATAGGTTA GGATAAAATATCCAAT 237	IV CCCCTATITGTTTATIT GGGGATAAACAAATAAA P_I Mbo II	ATTGAAAAG TAACTTTTTC 1 Hbh T	TTTGCTCAC AAACGAGTG	
Sfan I Sauge I Sauge I Fnu4H I Eco0109 I Fnu4H I Eco0109 I Fnu4H I Eco0109 I Hae III Hae III Hae III Hae III Sauge I II Hae III Hae II	II II II Hinp I TCAGGTGGCACTTTTCGGGGAATGTGCGCGGAA AGTCCACCGTGAAAGCCCCTTTACACGCGCGTT AGTCCACCGTGAAAGCCCCTTTACACGCGCTT AGTCCACCGTGAAAGCCCCTTTACACGCGCCTT AGTCCACCGTGAAAGCCCCTTTACACGCCCTTTACACGCCGCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACACGCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCCTTTACACGCCCTTTACACGCCCCTTTACACGCCCCCCTTTACACGCCCCTTTACACACGCCCTTTACACACGCCC	TGAGACAATAACCCTGATAAATGCTTCAATAAT ACTCTGTTATTGGGACTATTTACGAAGTTATTA ACTCTGTTATTGGGACTATTTACGAAGTTATTA 252	GGCA CCGT	

	57/	[′] 68		7L
2720		2800	2880	Fig.
iA I Mae III A W I NSPB II AIW I GCACGAGTGCTAAAAGCGGGG 27	656 2668 2674 2681 2690 2698 656 2674 2681 2698 2682 2699 2682 2699 2682 2699 2699 2699 1	GCACTITTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGA GCACTITTAAAGTTCCAAGACGATACACCGCGCCATAATAGGGCATAACTGCGGCCGGTTCT CGTGAAATTTCAAGACGATACACCGCGCCATAATAGGGCATAACTGCGGCCCGTTCT 41 2787 41 2749 2767 41 2748 2791 2791	RSQ I HPH I I DOOR I SCOT I HPH I STON I FOK I ATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATG 288 TATGTGATAAGAGTCTTACTGAACCAACTCATGAGTGGTCTTTTCGTAGAATGCCTAC 2827 2845 2856 2867 2876	FI
Sau3A I Mbo I Dpn I o II e TTAGTTGG	2646 2646 2646 2646 Xmn I	2 - 4 - 6 - 6 - 6 - 6 - 6 - 6 - 6 - 6 - 6	Fnu4H GCAACTCGGTCGCCGC CGTTGAGCCAGCGGCG	- 5

ı		58/68		7M
I 2960		3040	3120	Fig.
2850 Hae III Gdi I Eae I CCATAACATGCGGCCAACTTACTTCTGACAAC	918C1 CATICICITIAN ACCICACION CONTROL CACION CONTROL CACION CONTROL CACION CACI	HZ-UU-M MMM 2007 PH 2008 4+ €000	S to S	2

Hinp I Hha I Acccc 3200 Acccc 3200 3197 3197 3197 3197 3197 3197 3160 Acccc 3200 3197 3197	r _
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Fig. 7N
- 14 CC PU	
BS + U I BS + U I GAGGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	
Fok I HP I GGGGGTA 3163 3168 3163 3168 WHP I HP I	
SP I NSB I N	
A I U I NC I BCT CTTGATGAGGGCCTTGCTTCCCCCTTGATGAGGGCCCTTCCCCTTCCCCTTCCCCTTCCCCTTCCCCTTCCCCTTCCCC	1

60/68			7
3440 3440 3520		00	F.1.9
CAAGITIACICATATATACITITAGATIGAT GITCAAATGAGTATATATGAAATCTAACTA 3496 3496 3496 3512	I Sau3A I Mbo I Bstu I II Dpn I HinP I Bsty I Hha I	GAACTCTAGGAAAAAAGAC	
111 111 1GACAGTCTG 1GACAGTCTG 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Squ3A Mbo I Dpn I A I Mbo	GTTTCCTAGAA	
B		TAGAAAGAT	
Mn I I MS MS	9	TGACTCGCAGTCTGGGGC	
Sau 3A I Mbo I Dpn I CTAGCGACTCI 3362 3362 3362 3362 3362 348 I TTAAAACTTCA AATTTGAAGT		GTTTTCGTTCC, CAAAAGCAAGG	

	61/68	8	7P
	3680	3760 I	1 3840 Fig.
3590 3590 3591	CCAACTCTTTTTCCGAAG GGTTGAGAAAAGGCTTÇ	CCACCACTTCAAGAACTC GCACCACTTCAAGAACTC GGTGGTGAAGTTCTTGAG 0 41 ScrF	HPG BCN TAGTCGTGTCTTACCG ATTCAGCACAGAATGGC
3575 3576 3576 3576 3576	au3A I bo I I I Alu I ATCAAGAGCTA TAGTTCTCGAT 3658	651 Hqe CCGTAGTTAGG GCCATCATCC 374 u4H I	V I TGCCAGTGGCGA ACGGTCACCGCT
3564 3564 3565 3565 3565	3363 T+h11 HPD NSP DD N	I TGTA ACAT	Alwn I Bb III Bbv I ACCAGIGGCIGC
3557 3557 3557 3557	NSPB II LACCAGCGGTGGT ATGGTCGCCACCA	CCAAATACTGTC GGTTTATGACAG	Mae GCTAATCCTGTT CGATTAGGACAA
3533	111 II I THIII II TIGCAACAAAAAACCACCGC AACGTTTGTTTTTTGGTGGCG 3609	HinP I Hhd I CTGCCTTCAGCAGAGCGCAGATA GACCGAAGTCGTCTCGCGTCTAT 3699 3699	Mn! I GCACCGCCTACATACCTCGCTCT
	Thu 4H Bbv 1 GCTGC CGACG 3601 3601	MG 6 GTAA(CATTO	ACAT

	62	2/68		70
.	3920	4000	4080	Fig.
- - - - - - - - - -	GGTTGGACTCAAGACGATAATTCCGGCGCGCGGGCTGAACGGGGGTTCGTGCACACCCAAGCCTTG CCAACCTGAGTTCTGCTATCAATGCCGTCGCGTCG	HinP I Hha I Hae I I CGTGAGCATTGAGAAGCGCCA GCACTCGTAACTCTTTCGCGGT 3972 3973 3973 3973	Msp I Fnu4H I Hinp I BstN I EcoR II EcoR II EcoR II EcoR II BstN I CCTCCGGTAGGGGGGGAACGCCTGGTATCTTT CCTGCGTAGGGCCTTTTTTTTTT	

			63/68		7R
	4160		4240	I 4320	Fig.
4069	4069 NID IV GGCGGAGCCTATGGAA	CCCCCCCTCGGATACCTTT 4147 111 24 1	II TTCTTTCCTGCGTTATCCCC AAGAAAGGACGCAATAGGGG	Fle I Fnu4H I Bbv I HinP I Hhq I Hinf ACGACCGAGCGAGT TTGCTGGCTCG	
405	4046 4056 4056 4056 \$fan I Sfan I	AAACACTACGAGCAGTC 4129 NPH NSPH HGE_III NSPH	HGB I A11 I 16GCCTTTTGCTCACATG ACCGGAAACGAGTGTAC 4203 4216 4204 4216 4216 4216	Fnu4H I Bbv I Bbv I Fnu4H I TACCGCTCGCCGCAGCCG	291 4294 4294
	Mni I Hga I	TGGAGACTGAACTCGCAGCTA 4103 4115 4118 Hae III Scrf I Ecor II	166CCTTTTGC ACCGGAAACG 	TACCGCCTTTGAGTGAGCTGA	4277
4011 4018		CAAAGCGG CAAAGCGG Hae III	I CCTTT GGAAA	Hinf I TGATTCTGTGGATAACCGTAT	4242

```
64/68
                                             SCT
BCCT
BS + OR
IN
                                                        N
Ban
Gd÷
Ede
                                                                                           MSP
```

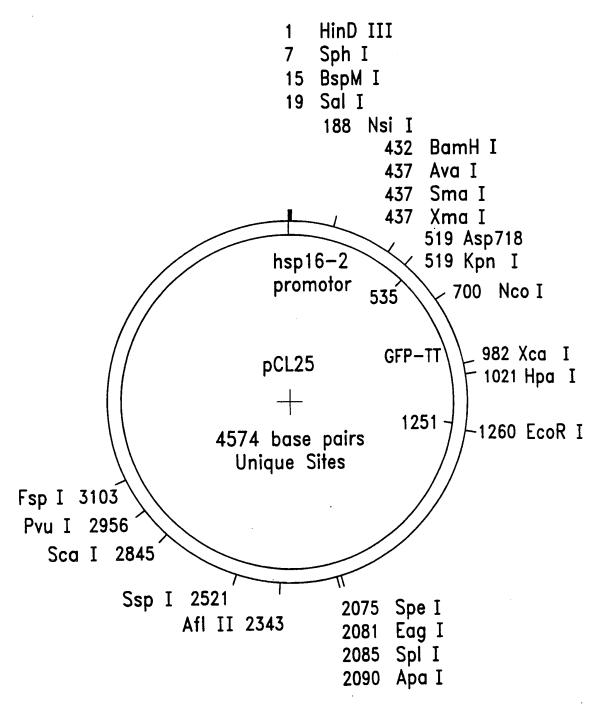


Fig. 8

aggaccagat gggcacgatg gacaagacgg agtcccagga acgiccagga agagacggaa acccaggaat gccaggagac gggccgccca ggaccaagag gaccaccagg agaggctggg aggatccgat ggaaagccag gatctccagg aggaaaggga acaaggacca aatggagctc caggactcag aggaatgcgc gcettcaagg atatggagga getgetggag aggaegggta ggatgaaaa gatgettate caggacagte tggaccacaa gccgaatact gcaaatgccc aggaagagaa ggagatgctg accaaatgga atgtcaatgc actgttgaga actcatgccc accaggacca gaatggagga atggagecae tggtgttcag ccaccagcae caactecaaa eecatatgga ttgaeggaa aagatgeega agatgtteaa aaeaeeeeae eaaeaggatg etteaeetgt gacttettgt aaatataaaa attteagaee agaaggaeea eeaggageee tggttaaact tggagcagga accgcttcca accgtgtgag acgtcaacaa atctaactga aaatttecag atgacectaa ctaeggegae gteeggegee ccaacticag agigagitic cgcaaattcc aattgtagac aaattcatga catttttcca Ittetaatte aetgataatt teaattattt agagteteae tgaagaeatg ctttateatt teaettitte tiggeaeteg tecaaaaate teaattaaaa attietggat catactictc agaattitit gittigtati iggittacca iggaigtiaa Highiggti tetetititig ecgetgette getttaeagt atcacaatga catcettatg tatatgcatt ettittattg titeetgata titattetel cagatggtat gtccattccc gagaaattcc aggaggtgga ttcccagatg tgagctggat gccgaaatcg attaaaatac aaaactgtat cacagggacc aactggaaga agetecacea aaggagagga gaggaaatca caagecagee gacaaccagg caddddcccc agagaccatt caaagttcc cacttggacc aggaaaagat cagacgtcat ggggaaaata atatctggaa attgtattt ctggagccac cacaaggac cagaaggac ggagagcag aaaaataaaa gacgaagtgc ggaaatcaac gacgatggag ggatatggag ggaccacgtg gctggaccag ggagctcgtg gtggaccac caggacttcc gggttgata aatattaaac statatagca atggaggat saagttteca statetttag gaactttett 60 66 72 78 $\frac{\infty}{2}$ 241 301 48 54, 08

Fig. 9

DCGPPGAPGSDGKPGSPGGKGDDGERPLGRPGPRGPPGEAGPEGPQGPTGRDAYPGQSG KFPGGGFPDGPFPNGGGPRGGNQCQCTVENSCPPGPAGPEGEEGPDGHDGQDGVPGFD GKDAEDVQNTPPTGCFTCPQGPLGPQGPNGAPGLRGMRGARGQPGRPGRDGNPGMPG **LGAGTASNRVRRQQYGGYGATGVQPPAPTPNPYGGYGASQPAPPEKFPDGIPNGGNQP** PQGEPGLQGYGGAAGEDGPEGPPGAPGLPGKDAEYCKCPGREGDAGRSARRHRKFQL MTLTTATSGAIVFSGATLLVSLFAAASLYSQVSNIWNELDAEIANFRSLTEDMWVDMVK

Fig. 10

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:
C12N 15/00, A01K 67/033, C07K 14/435,
14/47

(11) International Publication Number: WO 98/28971

(43) International Publication Date: 9 July 1998 (09.07.98)

(21) International Application Number: PCT/US97/23819

(22) International Filing Date: 31 December 1997 (31.12.97)

(30) Priority Data:

08/778,645 3 January 1997 (03.01.97) US 60/080,222 3 January 1997 (03.01.97) US

(71) Applicant (for all designated States except US): UNIVER-SITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LINK, Christopher [US/US]; 1837 Spruce Street, Boulder, CO 80302 (US).

(74) Agents: MONROY, Gladys, H. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

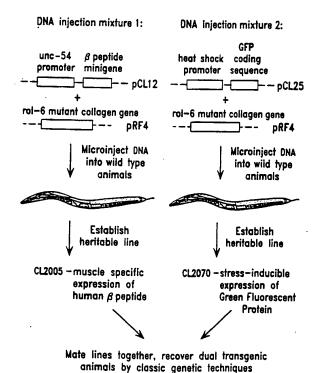
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
29 October 1998 (29.10.98)

(54) Title: TRANSGENIC ANIMALS FOR IN VIVO VISUALIZATION OF BETA-AMYLOID TOXICITY

(57) Abstract

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes in vivo. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, Caenorhabditis elegans for visualization of the toxic effects of β-amyloid accumulation in vivo.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	ТJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	ΙE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	ΚZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Lr. ational Application No PCT/US 97/23819

		PCT/US 97	7/23819
A. CLASS IPC 6	SIFICATION OF SUBJECT MATTER C12N15/00 A01K67/033 C07K	14/435 C07K14/47	
According	to International Patent Classification(IPC) or to both national cla	ssilication and IPC	
	S SEARCHED	Sancation and 17-9	
Minimum o	documentation searched (classification system followed by class	ification symbols)	
IPC 6	A01K C07K		
Documenta	ation searched other than minimum documentation to the extent (that such documents are included in the fields se	arched
Electronic	data base consulted during the international search (name of da	ta base and, where practical, search terms used)
			•
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT		
Category 3	Citation of document, with indication, where appropriate, of th	e relevant passages	Relevant to claim No.
Χ	CA 2 088 379 A (CANDIDO EDWARD	P M	1 4 6 7
	STRINGHAM EVE G (CA); JONES D	ONALD (CA))	1-4,6,7, 13-15,20
Υ	30 July 1994 see the whole document	ı	1-23
Y			1-23
	LINK, C.D.: "Expression of hubeta-amyloid peptide in transg	man enic	1-8, 10-15,
	Caenorhabditis elegans"		18-20,
	PROCEEDINGS OF THE NATIONAL AC. SCIENCES OF USA		22,23
	vol. 92, no. 20, 26 September 9368-9372, XP002065639 WASHINGTON US	1995, pages	
	cited in the application		ts. Name
X	see the whole document		11-15
		-/	
	ner documents are listed in the continuation of box C.	X Patent family members are listed in	annex.
	tegories of cited documents : Int defining the general state of the art which is not	"T" later document published after the inter	national filing date
conside	ered to be of particular relevance locument but published on or after the international	or priority date and not in conflict with t cited to understand the principle or the invention	ory underlying the
illing da documer "_	ate nt which may throw doubts on priority claims or	"X" document of particular relevance; the cl. cannot be considered novel or cannot in	pe considered to
citation	s cited to establish the publication date of another or other special reason (as specified)	"Y" document of particular relevance: the cla cannot be considered to involve an inv	aimed invention
otner m		document is combined with one or mor ments, such combination being obvious	e other such docu-
later the	nt published prior to the international filing date but an the priority date claimed	in the art. "&" document member of the same patent for	amily
ate of the a	ctual completion of theinternational search	Date of mailing of the international search	ch report
2	September 1998	1 1 09 1998	
ame and ma	ailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2	Authorized officer	
	NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040. Tx. 31 651 epo nl.		
	Fax: (+31-70) 340-3016	Chambonnet, F	

Jnt :tional Application No
PCT/US 97/23819

		PC1/US 9//23819
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	٠.
Category :	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LINK, C.D.: "Transgenic Caenorhabditis elegans as model system to study amyloid formation and toxicity" NEUROBIOLOGY OF AGING, vol. 15, no. suppl. 1, 29 July 1994, page S33 XP002065640 see the whole document	1,5-7
X	WO 95 11968 A (ATHENA NEUROSCIENCES INC; LILLY CO ELI (US)) 4 May 1995 see page 38, line 3 - page 39, line 22	11,12,19
Υ	WO 93 23533 A (TSI CORP) 25 November 1993 see page 27, paragraph 3 - page 28, paragraph 2	11-15
Y	CHALFIE M ET AL: "GREEN FLUORESCENT PROTEIN AS A MARKER FOR GENE EXPRESSION" SCIENCE, vol. 263, no. 5148, 11 February 1994, pages 802-805, XP002003599 see the whole document	1,2.6,7, 9,11, 13-17, 20,21
Y	MOSS, J. B. ET AL.: "Green fluorescent protein marks skeletal muscle in murine cell lines and zebrafish" GENE, vol. 173, no. 1, 1996, pages 89-98, XP002065641 AMSTERDAM NL see the whole document	1,2,9, 13,16,17
;		

· 4

Int. _tional application No. PCT/US 97./23819

B x I Observations where certain claim were found unsearchable (Continuation of item 1 of first sheet)
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)
This International Searching Authority found multiple inventions in this international application, as follows:
÷
see additional sheet
1. X As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Remark on Protest The additional search fees were accompanied by the applicant's protest. X No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-10 and partially 11-23

a transparent animal comprising:

(a) a toxic transgene;

- (b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo; and a method for providing such a dually transgenic transparent non-human animal comprising:
 - a) providing:
 - i) a first and a second non-human transparent animals;
 - ii) a first toxic transgene; and
 - iii) a second transgene, comprising an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo,
 - b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
 - c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic transparent non-human animal wherein said toxic transgene is expressed and induces the promotor so that the said reporter gene is also expressed and detectable in vivo:
 - 2. Claims: partially 11-23

a method to produce a dually transgenic non-human animal with a toxic transgene and a reporter are expressed comprising:

- a) providing:
- i) a first and a second non-human animal:
- ii) a first toxic transgene: and
- iii) a second transgene, comprising a reporter,
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic non-human animal wherein said toxic transgene and said reporter are expressed

in the cases non relevant of the first described subject, that means where it is not a transparent animal comprising :

(a) a toxic transgene;

(b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo.

Information on patent family members

Ir. ational Application No PCT/US 97/23819

Patent document cited in search repor	t	Publication date		atent family nember(s)	. Publication date
CA 2088379 A		30-07-1994	NONE	· · · · · · · · · · · · · · · · · · ·	,
WO 9511968	- 	04-05-1995	US	5604102 A	18-02-1997
			US	5612486 A	18-03-1997
			ĀŪ	8079894 A	22-05-1995
			AU	8080994 A	22-05-1995
			CA	2174429 A	04-05-1995
			CA	2174632 A	04-05-1995
			EP	0736106 A	09-10-1996
			EP	0730643 A	11-09-1996
			JP	9508196 T	19-08-1997
			JP	9507746 T	12-08-1997
			WO	9511994 A	04-05-1995
WO 9323533	<u>-</u> -	25-11-1993	AU	4240293 A	13-12-1993